(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 February 2002 (07.02.2002)

PCT

(10) International Publication Number
WO 02/10382 A2

(51) International Patent Classification[7]: C12N 15/12, 15/11, 9/00, C07K 14/47, C12Q 1/68, G01N 33/577, A61K 31/713

(21) International Application Number: PCT/EP01/08309

(22) International Filing Date: 18 July 2001 (18.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/221,513 28 July 2000 (28.07.2000) US

(71) Applicant and
(72) Inventor: **WISSENBACH, Ulrich** [DE/DE]; Institut für Pharmakologie und Toxikologie der Uni, versität des Saarlandes, 66421 Homburg (DE).

(74) Agent: **HUBER, Bernard**; Huber & Schüssler, Truderinger Str. 246, 81825 München (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/10382 A2

(54) Title: TRP8, TRP9 AND TRP10, NOVEL MARKERS FOR CANCER

(57) Abstract: The present invention relates to gene expression in normal cells and cells of malignant tumors and particularly to novel markers associated with cancer, Trp8, Trp9 and Trp10, and the genes encoding Trp8, Trp9 and Trp10. Also provided are vectors, host cells, antibodies, and recombinant methods for producing these human proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating a tumor.

Trp8, Trp9 and Trp10, novel markers for cancer

FIELD OF THE INVENTION

The present invention relates to gene expression in normal cells and cells of malignant tumors and particularly to novel markers associated with cancer, Trp8, Trp9 and Trp10, and the genes encoding Trp8, Trp9 and Trp10

BACKROUND OF THE TECHNOLOGY

Prostate cancer is one of the most common diseases of older men world wide. Diagnosis and monitoring of prostate cancer is difficult because of the heterogeneity of the disease. For diagnosis different grades of malignancy can be distinguished according to the Gleason-Score Diagnosis. For this diagnosis a prostate tissue sample is taken from the patient by biopsy and the morphology of the tissue is investigated. However, this approach only yields subjective results depending on the experience of the pathologist. For confirmation of these results and for obtaining an early diagnosis an additional diagnostic method can be applied which is based on the detection of a prostate specific antigen (PSA). PSA is assayed in serum samples, blood samples etc. using an anti-PSA-antibody. However, since in principle PSA is also expressed in normal prostate tissue there is a requirement for the definition of a threshold value (about 4 ng/ml PSA) in order to be able to distinguish between normal and malign prostate tissue. Unfortunately, this diagnostic method is quite insensitive and often yields false-positive results. Moreover, by using this diagnostic method any conclusions as regards the grade of malignancy, the progression of the tumor and its potential for metastasizing cannot be drawn. Thus, the use of molecular markers would be helpful to distinguish benign from malign tissue and for grading and staging prostate carcinoma, particularly for patients with metastasizing prostate cancer having a very bad prognosis.

The above discussed limitations and failings of the prior art to provide meaningful specific markers which correlate with the presence of prostate tumors, in particular metastasizing tumors, has created a need for markers which can be used diagnostically, prognostically and therapeutically over the course of this disease. The present invention fulfils such a need by the provision of Tpr8, Trp9 and Trp10 and the genes encoding Trp8, Trp9 and Trp10: The genes encoding Trp8 and Trp10 are expressed in prostate carcinoma and prostatic metastasis, but

not in normal prostate, benign hyperplasia (BHP) and intraepithelial prostatic neoplasia (PIN). Furthermore, expression of Trp10 transcripts is detectable in carcinoma but not in healthy tissue of the lung, the prostate, the placenta and in melanoma.

SUMMARY OF THE INVENTION

The present invention is based on the isolation of genes encoding novel markers associated witha cancer, Trp8, Trp9 and Trp10. The new calcium channel proteins Trp8, Trp9 and Trp10 are members of the trp (transient receptor potential) - family, isolated from human placenta (Trp8a and Trp8b) and humane prostate (Trp9, Trp10a and Trp10b). Trp proteins belong to a steadily growing family of Ca^{2+} selective and non selective ion channels. In the recent years seven Trp proteins (trp1 - trp7) have been identified and suggested to be involved in cation entry, receptor operated calcium entry and pheromone sensory signaling. Structurally related to the trp proteins are the vanilloid receptor (VR1) and the vanilloid like receptor (VRL-1) both involved in nociception triggered by heat. Furthermore, two calcium permeable channels were identified in rat small intestine (CaT1) and rabbit kidney (ECaC). These distantly related channels are suggested to be involved in the uptake of calcium ions from the lumen of the small intestine (CaT1) or in the reuptake of calcium ions in the distal tubule of the kidney (ECaC). Common features or the Trp and related channels are a proposed structure comprising six transmembrane domains including several conserved amino acid motifs. In the present invention the cloning and expression of a CaT1 like calcium channel (Trp8) from human placenta as well as Trp9 and Trp10 (two variants, Trp10a and Trp10b) is described. Two polymorphic variants of the Trp8 cDNA were isolated from placenta (Trp8a and Trp8b). Transient expression of the Trp8b cDNA in HEK (human embryonic kidney) cells results in cytosolic calcium overload implicating that the Trp8 channel is constitutive open in the expression system. Trp8 induces highly calcium selective inward currents in HEK cells. The C -terminus of the Trp8 protein binds calmodulin in a calcium dependent manner. The Trp9 channel is expressed in trophoblasts and syncytiotrophoblasts of placenta and in pancreatic acinar cells. Furthermore, the Trp8 channel is expressed in prostatic carcinoma and prostatic metastases, but not in normal tissue of the prostate. No expression of Trp8 transcripts is detectable in benign prostatic hyperplasia (BPH) or prostatic intraepithelial neoplasia (PIN). Therefore, the Trp8 channel is exclusively expressed in malign prostatic tissues and serves as molecular marker for prostate cancer. From the experimental results it is also apparent that the

modulation of Trp8 and/or Trp10, e.g. the inhibition of expression or activity , is of therapeutic interest, e.g. for the prevention of tumor progression.

The present invention, thus, provides a Trp8, Trp9 and Trp10 protein, respectively, as well as nucleic acid molecule encoding the protein and, moreover, an antisense RNA, a ribozyme and an inhibitor, which allow to inhibit the expression or the activity of Trp8, Trp9 and/or Trp10.

In one embodiment, the present invention provides a diagnostic method for detecting a prostate cancer or endometrial cancer (cancer of the uterus) associated with Trp8 or Trp10 in a tissue of a subject, comprising contacting a sample containing Trp8 and/or Trp10 encoding mRNA with a reagent which detects Trp8 and/or Trp10 or the corresponding mRNA.

In a further embodiment, the present invention provides a diagnostic method for detecting a melanoma, chorion carcinoma, cancer of the lung and of the prostate in a tissue of a subject, comprising contacting a sample with a reagent which detects Trp10a and/or Trp10b antisense transcripts or Trp10a and/or Trp10b related antisense transcripts.

In another embodiment, the present invention provides a method of treating a prostate tumor, carcinoma of the lung, carcinoma of the placenta (chorion carcinoma) or melanoma associated with Trp8 and/or Trp10, comprising administering to a subject with such an disorder a therapeutically effect amount of a reagent which modulates, e.g. inhibits, expression of Trp8 and/or Trp10 or the activity of the protein, e.g. the above described compounds.

Finally, the present invention provides a method of gene therapy comprising introducing into cells of a subject an expression vector comprising a nucleotide sequence encoding the above mentioned antisense RNA or ribozyme, in operable linkage with a promoter.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1: A, phylogenetic relationship of trp and related proteins. B, hydropathy plot of the Trp8 protein sequence according to Kyte and Doolittle. C, alignment of Trp8a/b to the epithelial calcium channels ECaC (from rabbit) and Vr1 (from rat). Putative transmembrane domains are underlined.

Figure 2: A, polymorphism of the Trp8 gene. The polymorphic variants Trp8a and Trp8b differ in five base pairs resulting in three amino acid exchanges in the derived protein sequences. Specific primers were derived from the Trp8 gene as indicated by arrows. B, the Trp8a and Trp8b genes are distinguishable by a single restriction site. Genomic fragments of the Trp8 gene can be amplified using specific primers (shown in A). The genomic fragment of the Trp8b gene contains an additional site of the restriction enzyme BSP1286I (B). C, the Trp8 gene is located on chromosome 7. D, genotyping of eleven human subjects. A 458 bp genomic fragment of the Trp8 gene was amplified using specific primers (shown in A) and restricted with BSP1286I. The resulting fragments were analyzed by PAGE electrophoresis.

Figure 3: The Trp8b protein is a calcium selective ion channel. A, representative trace of a pdiTrp8b transfected HEK 293 cell. Trp8b mediated currents are activated by voltage ramps (-100 mV - +100 mV) of 100 msec at -40 mV or +70 mV holding potential. 1, Trp8b currents in the presence at 2mm $[Ca^{2+}]_o$;2, effect of solution switch alone 3, switch to nominal zero calcium solution. B, Trp8b currents in the presence of zero divalent cations. C, current voltage relationship of the currents shown in A. Inset, leak subtracted current. D, current voltage relationship of the current shown in B. E, statistics of representative experiments. Black: Trp8 transfected cells, gray: control cells. Columns from left to right: Trp8 currents at - 40 mV (n =12) and + 70 mV holding potential (n = 12). Trp8 currents in standard bath solution including 120 mM NMDG without sodium (n = 7) and with nominal zero calcium ions (n = 8) or in the presence of 1mM EGTA with zero divalent cations (n = 6). F, representative changes in $[Ca^{2+}]_i$ in Trp8b transfected HEK cells (gray) and controls (black) in the presence or absence of 1mM $[Ca^{2+}]_o$. Inset, relative increase of cytosolic calcium concentration of Trp8b transfected HEK cells, before and after readdition of 1 mM $[Ca^{2+}]_o$ in comparison to control cells.

Figure 4: The C-terminal region of the Trp8 protein binds calmodulin. A, N- and C-terminal fragments of the Trp8 protein used for calmodulin binding studies. B, the Trp8 protein and a truncated Trp8 protein which was in vitro translated after MunI cut of the cDNA, which lacks the C-terminal 32 amino acid residues, were in vitro translated in the presence of ^{35}S-methionine and incubated with calmodulin coupled agarose beads in the presence of 1 mM Ca^{2+} or 2 mM EGTA. C, calmodulin binding to N- and C-terminal fragments of the Trp8protein in the presence of Ca^{2+} (1 mM) or EGTA (2 mM)

Figure 5: Expression pattern of the Trp8 cDNA. A, Northern blots (left panels, Clontech, Palo Alto) were hybridized using a 348 bp NcoI/BamHT fragment of the Trp9 cDNA. The probe hybridizes to mRNA species isolated from the commercial blot, but not to mRNA species isolated from benign prostate hyperplasia (right panel, mRNA isolated from 20 human subjects with benign prostate hyperplasia). B,C, in situ hybridization with biotinylated Trp8 specific oligonucleotides on slides of human tissues. Left column antisense probes, right column sense probes. D, antinsense probes.

Figure 6: Differential expression of Trp8 cDNA in human prostate. A - F, in situ hybridization with prostatic tissues. A, normal prostate, B, primary carcinoma, C, benign hyperplasia, D, rezidive carcinoma, E, prostatic intraepithelial neoplasia, F, lymphnode metastasis of the prostata.

Figure 7: Trp8a cDNA sequence and derived amino acid sequence

Figure 8: A, Trp8b cDNA sequence and derived amino acid sequence
B, cDNA sequence of splice variant 1 (12B1)
C, cDNA sequence of splice variant 2 (17-3)
D, cDNA sequence of splice variant 3 (23A3)
E, cDNA sequence of splice variant 4 (23C3)

Figure 9: A, Trp9 cDNA sequence and derived amino acid sequence B, cDNA sequence of splice variant 15 and derived amino acid sequence.

Figure 10: A, cDNA sequence of Trp10a and derived amino acid sequence, B, cDNA fragment of Trp10a and derived amino acid sequence.

Figure 11: cDNA sequence of Trp10b and derived amino acid sequence.

Figure 12: Expression of Trp8 mRNA in human endometrial cancer or cancer of the uterus. A - D, in situ hybridization with slides of endometrial cancer hybridized with Trp8 antisense (left column) or sense probes as controls (right column). E - F, Trp8 antisense probes hybridized to slides of normal endometrium. It can be clearly seen no hybridization occurs with normal endometrial tissue.

Figure 13: Expression of human Trp9 and Trp10 genes
Northern blots were hybridized using Trp9 (upper panel) or Trp10 (lower panel) specific probes. Expression of the Trp9 cDNA is detectable in many tissues including human prostate and colon as well as in benign prostatic hyperplasia. Expression of Trp10 cDNA is detectable in human prostate of a commercial northern blot (Clontech, right side). This Northern blot contains prostatic tissue collected from 15 human subjects in the range of 14 - 60 years of age. No expression of Trp10 cDNA was detectable in benign prostatic hyperplasia (left side).

Figure 14: Expression of Trp10 transcripts and Trp10-antisense transcripts in human prostate cancer and metastasis of a melanoma. In situ hybridizations of slides hybridized with Trp10-antisense (A-E, K-N) and Trp10 related sense probes (F-J, P-R). It can clearly be seen that both probes detect the same cancer cells indicating that these cancer cells express Trp10 transcripts as well as Trp10-antisense transcripts. S, no Trp10 expression is detectable in benign hyperplasia of the prostate (BPH). O and T, show expression of Trp10 transcripts (O) and Trp10-antisense transcripts (T) in a metastasis of a melanoma in human lung. Melanoma cancer cells express both Trp10 transcripts and Trp10-antisense transcripts.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to an isolated nucleic acid molecule encoding the human prostate carcinoma associated protein Trp8a, Trp8b, Trp9, Trp10a or Trp10b or a protein exhibiting biological properties of Trp8a, Trp8b, Trp9, Trp10a or Trp10b and being selected from the group consisting of

(a) a nucleic acid molecule encoding a protein that comprises the amino acid sequence depicted in Figure 7, 8A, 9,10 or 11;

(b) a nucleic acid molecule comprising the nucleotide sequence depicted in Figure 7, 8A, 9,10, or 11;

(c) a nucleic acid molecule included in DSMZ Deposit no. DSM 13579 (deposit date: 28 June 2000), DSM 13580 (deposit date: 28 June 2000), DSM 13584 (deposit date: 5 July 2000), DSM 13581 (deposit date: 28 June 2000) or DSM ....(deposit date:....);

(d) a nucleic acid molecule with hybridizes to a nucleic acid molecule specified in (a) to (c)

(e) a nucleic acid molecule the nucleic acid sequence of which deviates from the nucleic sequences specified in (a) to (d) due to the degeneration of the genetic code; and

(f) a nucleic acid molecule, which represents a fragment, derivative or allelic variation of a nucleic acid sequence specified in (a) to (e).

As used herein, a protein exhibiting biological properties of Trp8a, Trp8b, Trp9,Trp10a or Trp10b is understood to be a protein having at least one of the activities as illustrated in the Examples, below.

As used herein, the term „isolated nucleic acid molecule„ includes nucleic acid molecules substantially free of other nucleic acids, proteins, lipids, carbohydrates or other materials with which it is naturally associated.

In a first embodiment, the invention provides an isolated nucleic acid molecule encoding the human prostate carcinoma associated protein Trp8a, Trp8b, Trp9,Trp10a or Trp10b comprising the amino acid sequence depicted in Figure 7, 8A, 9,10 or 11. The present invention also provides a nucleic acid molecule comprising the nucleotide sequence depicted in Figure 7, 8A, 9,10 or 11.

The present invention provides not only the generated nucleotide sequence identified in Figure 7, 8A, 9,10 or 11, respectively and the predicted translated amino acid sequence, respectively, but also plasmid DNA containing a Trp8a cDNA deposited with the DSMZ, under DSM 13579, a Trp8b cDNA deposited with the DSMZ, under DSM 13580, a Trp9 cDNA deposited with the DSMZ, under DSM 13584, a Trp10a cDNA deposited with the DSMZ, under DSM 13581, and a Trp10b cDNA deposited with the DSMZ, under DSM...., respectively. The nucleotide sequence of each deposited Trp-clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by each deposited clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited Trp-encoding DNA, collecting the protein, and determining its sequence.

The nucleic acid molecules of the invention can be both DNA and RNA molecules. Suitable DNA molecules are, for example, genomic or cDNA molecules. It is understood that all

nucleic acid molecules encoding all or a portion of Trp8a, Trp8b, Trp9,Trp10a or Trp10b are also included, as long as they encode a polypeptide with biological activity. The nucleic acid molecules of the invention an be isolated from natural sources or can be synthesized according to know methods.

The present invention also provides nucleic acid molecules which hybridize to the above nucleic acid molecules. As used herein, the term „hybridize„ has the meaning of hybridization under conventional hybridization conditions, preferably under stringent conditions as described, for example, in Sambrook et al., Molecular Cloning, A Laboratory Manual 2nd edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. Also contemplated are nucleic acid molecules that hybridize to the Trp nucleic acid molecules at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency), salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°Cin a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 9.2M NaH_2PO_4; 0.02M EDTA, pH7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA, following by washes at 50°C with 1 X SSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC). Variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Nucleic acid molecules that hybridize to the molecules of the invention can be isolated, e.g., from genomic or cDNA libraries that were produced from human cell lines or tissues. In order to identify and isolate such nucleic acid molecules the molecules of the invention or parts of these molecules or the reverse complements of these molecules can be used, for example by means of hybridization according to conventional methods (see, e.g., Sambrook et al., supra). As a hybridization probe nucleic acid molecules can be used, for example, that have exactly or basically the nucleotide sequence depicted in Figure 7, 8A, 9,10 or 11, respectively, or parts of these sequences. The fragments used as hybridization probe can be synthetic

fragments that were produced by means of conventional synthetic methods and the sequence of which basically corresponds to the sequence of a nucleic acid molecule of the invention.

The nucleic acid molecules of the present invention also include molecules with sequences that are degenerate as a result of the genetic code.

In a further embodiment, the present invention provides nucleic acid molecules which comprise fragments, derivatives and allelic variants of the nucleic acid molecules described above encoding a protein of the invention. „Fragments„ are understood to be parts of the nucleic acid molecules that are long enough to encode one of the described proteins. These fragments comprise nucleic acid molecules specifically hybridizing to transcripts of the nucleic acid molecules of the invention. These nucleic acid molecules can be used, for example, as probes or primers in the diagnostic assay and/or kit described below and, preferably, are oligonucleotides having a length of at least 10, in particular of at least 15 and particularly preferred of at least 50 nucleotides. The nucleic acid molecules and oligonucleotides of the invention can also be used, for example, as primers for a PCR reaction. Examples of particular useful probes (primers) are shown in Tables 1 and 2.

Table 1

Trp8 probes used for in situ hybridization:

Probes (antisense)

1.) 5' TCCGCTGCCGGTTGAGATCTTGCC 3'

2.) 5' CTTGCTCCATAGGCAGAGAATTAG 3'

3.) 5' ATCCTCAGAGCCCCGGGTGTGGAA3'

Controls (sense)

1.) 5' GGCAAGATCTCAACCGGCAGCGGA 3'

2.) 5' CTAATTCTCTGCCTATGGAGCAAG 3'

3.) 5' TTCCACACCCGGGGCTCTGAGGAT 3'

Tabelle 2

Trp10 probes used for the in situ hybridizations shown in Figure 14:

Probes (antisense)

1.) 5' GCTTCCACCCCAAGCTTCACAGGAATAGA 3' (Figure 14 A, 14B)

2.) 5' GGCGATGAAATGCTGGTCTGTGGC 3' (Figure 14C, 14D, 14N, 14S, 14O)

3.) 5' ATCTTCCAGTTCTTGGTGTCTCGG 3' (Figure 14E, 14K)

4.) 5' GCTGCAGTACTCCTGCACCAGGAA 3' (Figure 14L, 14M)

Probes (sense)

1.) 5' TCTATTCCTGTGAAGCTTGGGGTGGAAGC 3' (Figure 14F, 14G)

2.) 5' GCCACAGACCAGCATTTCATCGCC 3' (Figure 14H, 14I, 14T)

3.) 5' CCGAGACACCAAGAACTGGAAGAT 3' (Figure 14J, 14P)

4.) 5' TTCCTGGTGCAGGAGTACTGCAGC 3' (Figure 14Q, 14R)

The term „derivative„ in this context means that the sequences of these molecules differ from the sequences of the nucleic acid molecules described above at one or several positions but have a high level of homology to these sequences. Homology hereby means a sequence identity of at least 40%, in particular an identity of at least 60%, preferably of more than 80% and particularly preferred of more than 90%. These proteins encoded by the nucleic acid molecules have a sequence identity to the amino acid sequence depicted in Figure 7, 8A, 9, 10 and 11, respectively, of at least 80%, preferably of 85% and particularly preferred of more than 90%, 97% and 99%. The deviations to the above-described nucleic acid molecules may have been produced by deletion, substitution, insertion or recombination. The definition of the derivatives also includes splice variants, e.g. the splice variants shown in Figures 8B to 8E and 9B.

The nucleic acid molecules that are homologous to the above-described molecules and that represent derivatives of these molecules usually are variations of these molecules that represent modifications having the same biological function. They can be naturally occurring variations, for example sequences from other organisms, or mutations that can either occur naturally or that have been introduced by specific mutagenesis. Furthermore the variations can be synthetically produced sequences. The allelic variants can be either naturally occurring variants or synthetically produced variants or variants produced by recombinant DNA processes.

Generally, by means of conventional molecular biological processes it is possible (see, e.g., Sambrook et al., supra) to introduce different mutations into the nucleic acid molecules of the invention. As a result Trp proteins or Trp related proteins with possibly modified biological properties are synthesized. One possibility is the production of deletion mutants in which nucleic acid molecules are produced by continuous deletions from the 5'- or 3'-terminal of the coding DNA sequence and that lead to the synthesis of proteins that are shortened accordingly. Another possibility is the introduction of single-point mutation at positions where a modification of the amino aid sequence influences, e.g., the ion channel properties or the regulations of the trp-ion channel. By this method muteins can be produced, for example, that possess a modified ion conducting pore, a modified K_m-value or that are no longer subject to the regulation mechanisms that normally exist in the cell, e.g. with regard to allosteric regulation or covalent modification. Such muteins might also be valuable as therapeutically useful antagonists of Trp8a, Trp8b, Trp9,Trp10a or Trp10b, respectively.

For the manipulation in prokaryotic cells by means of genetic engineering the nucleic acid molecules of the invention or parts of these molecules can be introduced into plasmids allowing a mutagenesis or a modification of a sequence by recombination of DNA sequences. By means of conventional methods (cf. Sambrook et al., supra) bases can be exchanged and natural or synthetic sequences can be added. In order to link the DNA fragments with each other adapters or linkers can be added to the fragments. Furthermore, manipulations can be performed that provide suitable cleavage sites or that remove superfluous DNA or cleavage sites. If insertions, deletions or substitutions are possible, in vitro mutagenesis, primer repair, restriction or ligation can be performed. As analysis method usually sequence analysis, restriction analysis and other biochemical or molecular biological methods are used.

The proteins encoded by the various variants of the nucleic acid molecules of the invention show certain common characteristics, such as ion channel activity, molecular weight, immunological reactivity or conformation or physical properties like the electrophoretical mobilty, chromatographic behavior, sedimentation coefficients, solubility, spectroscopic properties, stability; pH optimum, temperature optimum.

The invention furthermore relates to vectors containing the nucleic acid molecules of the invention. Preferably, they are plasmids, cosmids, viruses, bacteriophages and other vectors

usually used in the field of genetic engineering. Vectors suitable for use in the present invention include, but are not limited to the T7-based expression vector for expression in mammalian cells and baculovirus-derived vectors for expression in insect cells. Preferably, the nucleic acid molecule of the invention is operatively linked to the regulatory elements in the recombinant vector of the invention that guarantee the transcription and synthesis of an RNA in prokryotic and/or eukaryotic cells that can be translated. The nucleotide sequence to be transcribed can be operably linked to a promoter like a T7, metallothionein I or polyhedrin promoter.

In a further embodiment, the present invention relates to recombinant host cells transiently or stable containing the nucleic acid molecules or vectors or the invention. A host cell is understood to be an organism that is capable to take up *in vitro* recombinant DNA and, if the case may be, to synthesize the proteins encoded by the nucleic acid molecules of the invention. Preferably, these cells are prokaryotic or eukaryotic cells, for example mammalian cells, bacterial cells, insect cells or yeast cells. The host cells of the invention are preferably characterized by the fact that the introduced nucleic acid molecule of the invention either is heterologous with regard to the transformed cell, i.e. that it does not naturally occur in these cells, or is localized at a place in the genome different from that of the corresponding naturally occurring sequence.

A further embodiment of the invention relates to isolated proteins exhibiting biological properties of the human prostate carcinoma associated protein Trp8a, Trp8b, Trp9,Trp10a or Trp10b and being encoded by the nucleic acid molecules of the invention, as well as to methods for their production, whereby, e.g., a host cell of the invention is cultivated under conditions allowing the synthesis of the protein and the protein is subsequently isolated from the cultivated cells and/or the culture medium. Isolation and purification of the recombinantly produced proteins may be carried out by conventional means including preparative chromatography and affinity and immunological separations involving affinity with an anti-Trp8a-, anti-Trp8b-, anti-Trp9-,anti-Trp10a- or anti-Trp10b-antibody, respectively.

As used herein, the term „isolated protein„ includes proteins substantially free of other proteins, nucleic acids, lipids, carbohydrates or other materials with which it is naturally associated. Such proteins however not only comprise recombinantly produced proteins but include isolated naturally occurring proteins, synthetically produced proteins, or proteins

produced by a combination of these methods. Means for preparing such proteins are well understood in the art. The Trp proteins are preferably in a substantially purified form. A recombinantly produced version of a human prostate carcinoma associated protein Trp8a, Trp8b, Trp9,Trp10a or Trp10b protein, including the secreted protein, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67; 31-40 (1988).

In a further preferred embodiment, the present invention relates to an antisense RNA sequence characterised that it is complementary to an mRNA transcribed from a nucleic acid molecule of the present invention or a part thereof and can selectively bind to said mRNA, said sequence being capable of inhibiting the synthesis of the protein encoded by said nucleic acid molecules, and a ribozyme characterised in that it is complementary to an mRNA transcribed from a nucleic acid molecule of the present invention or a part thereof and can selectively bind to and cleave said mRNA, thus inhibiting the synthesis of the proteins encoded by said nucleic acid molecules. Riboyzmes which are composed of a single RNA chain are RNA enzymes, i.e. catalytic RNAs, which can intermolecularly cleave a target RNA, for example the mRNA transcribed from one of the Trp genes. It is now possible to construct ribozymes which are able to cleave the target RNA at a specific site by following the strategies described in the literature. (see, e.g., Tanner et al., in: Antisense Research and Applications, CRC Press Inc. (1993), 415-426). The two main requirements for such ribozymes are the catalytic domain and regions which are complementary to the target RNA and which allow them to bind to its substrate, which is a prerequisite for cleavage. Said complementary sequences, i.e., the antisense RNA or ribozyme, are useful for repression of Trp8a-, Trp8b, Trp9-,Trp10a- and Trp10b-expression, respectively, i.e. in the case of the treatment of a prostate cancer or endometrial cancer (carcinoma of the uterus). Preferably, the antisense RNA and ribozyme of the invention are complementary to the coding region. The person skilled in the art provided with the sequences of the nucleic acid molecules of the present invention will be in a position to produce and utilise the above described antisense RNAs or ribozymes. The region of the antisense RNA and ribozyme, respectively, which shows complementarity to the mRNA transcribed from the nucleic acid molecules of the present invention preferably has a length of at least 10, in particular of at least 15 and particularly preferred of at least 50 nucleotides.

In still a further embodiment, the present invention relates to inhibitors of Trp8a, Trp8b, Trp9, Trp10a and Trp10b, respectively, which fulfill a similar purpose as the antisense RNAs or

ribozymes mentioned above, i.e. reduction or elimination of biologically active Trp8a, Trp8b, Trp9, Trp10a or Trp10b molecules. Such inhibitors can be, for instance, structural analogues of the corresponding protein that act as antagonists. In addition, such inhibitors comprise molecules identified by the use of the recombinantly produced proteins, e.g. the recombinantly produces protein can be used to screen for and identify inhibitors, for example, by exploiting the capability of potential inhibitors to bind to the protein under appropriate conditions. The inhibitors can, for example, be identified by preparing a test mixture wherein the inhibitor candidate is incubated with Trp8a, Trp8b, Trp9, Trp10a or Trp10b, respectively, under appropriate conditions that allow Trp8a, Trp8b, Trp9, Trp10a or Trp10b to be in a native conformation. Such an in vitro test system can be established according to methods well known in the art. Inhibitors can be identified, for example, by first screening for either synthetic or naturally occurring molecules that bind to the recombinantly produced Trp protein and then, in a second step, by testing those selected molecules in cellular assays for inhibition of the Trp protein, as reflected by inhibition of at least one of the biological activities as described in the examples, below. Such screening for molecules that bind Trp8a, Trp8b, Trp9, Trp10a or Trp10b could easily performed on a large scale, e.g. by screening candidate molecules from libraries of synthetic and/or natural molecules. Such an inhibitor is, e.g., a synthetic organic chemical, a natural fermentation product, a substance extracted from a microorganism, plant or animal, or a peptide. Additional examples of inhibitors are specific antibodies, preferably monoclonal antibodies. Moreover, the nucleic sequences of the invention and the encoded proteins can be used to identify further factors involved in tumor development and progression. In this context it should be emphasized that the modulation of the calcium channel of a member of the trp family can result in the stimulation of the immune response of T lymphocytes leading to proliferation of the T lymphocytes. The proteins of the invention can, e.g., be used to identify further (unrelated) proteins which are associated with the tumor using screening methods based on protein/protein interactions, e.g. the two-hybrid-system Fields, S. and Song, O. (1989) Nature (340): 245-246.

The present invention also provides a method for diagnosing a prostate carcinoma which comprises contacting a target sample suspected to contain the protein Trp8a, Trp8b, Trp10a and/or Trp10b or the Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA with a reagent which reacts with Trp8a, Trp8b, Trp10a and/or Trp10b or the Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA and detecting Trp8a, Trp8b, Trp10a and/or Trp10b or Trp8a, Trp8b, Trp10a and/or Trp 10b encoding mRNA.

It has been found that carcinoma cells of placenta (chorion carcinoma), lung and prostate express Trp10 transcripts as well as Trp10 antisense transcripts and transcripts being in part complementary to Trp10 antisense transcripts. Accordingly, the present invention also provides a method for diagnosing a melanoma, chorion carcinoma, cancer of the lung and of the prostate in a tissue of a subject, comprising contacting a sample with a reagent which detects Trp10a and/or Trp10b antisense RNA.

When the target is mRNA (or antisense RNA), the reagent is typically a nucleic acid probe or a primer for PCR. The person skilled in the art is in a position to design suitable nucleic acids probes based on the information as regards the nucleotide sequence of Trp8a, Trp8b, Trp10a or Trp10b as depicted in figure 7, 8a, 10 and 11, respectively, or tables 1 and 2, above. When the target is the protein, the reagent is typically an antibody probe. The term „antibody", preferably, relates to antibodies which consist essentially of pooled monoclonal antibodies with different epitopic specifities, as well as distinct monoclonal antibody preparations. Monoclonal antibodies are made from an antigen containing fragments of the proteins of the invention by methods well known to those skilled in the art (see, e.g., Köhler et al., Nature 256 (1975), 495). As used herein, the term „antibody" (Ab) or „monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab') 2 fragments) which are capable of specifically binding to protein. Fab and f(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24: 316-325 (1983)). Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimerical, single chain, and humanized antibodies. The target cellular component, i.e. Trp8a, Trp8b, Trp10a and/or Trp10b or Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA or Trp10a/b antisense transcripts, e.g., in biological fluids or tissues, may be detected directly in situ, e.g. by in situ hybridization (e.g., according to the examples, below) or it may be isolated from other cell components by common methods known to those skilled in the art before contacting with a probe. Detection methods include Northern blot analysis, RNase protection, in situ methods, e.g. in situ hybridization, in vitro amplification methods (PCR, LCR, QRNA replicase or RNA-transcription/amplification (TAS, 3SR), reverse dot blot disclosed in EP-B1 O 237 362)), immunoassays, Western blot and other detection assays that are known to those skilled in the art.

Products obtained by in vitro amplification can be detected according to established methods, e.g. by separating the products on agarose gels and by subsequent staining with ethidium bromide. Alternatively, the amplified products can be detected by using labeled primers for amplification or labeled dNTPs.

The probes can be detectable labeled, for example, with a radioisotope, a bioluminescent, compound, a chemiluminescent compound, a fluorescent compound, a metal chelate, or an enzyme.

Expression of Trp8a, Trp8b, Trp10a and Trp10b, respectively, in tissues can be studied with classical immunohistological methods (Jalkanen et al., J. Cell. Biol. 101 (1985), 976-985; Jalkanen et al., J. Cell. Biol. 105 (1987), 3087-3096; Sobol et al. Clin. Immunpathol. 24 (1982), 139-144; Sobol et al., Cancer 65 (1985), 2005-2010). Other antibody based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (^{125}I, ^{121}I), carbon (^{14}C), sulfur (^{35}S), tritium (^{3}H), indium (^{112}In), and technetium rhodamine, and biotin. In addition to assaying Trp8a, Trp8b, Trp 10a or Trp10b levels in a biological sample, the protein can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma. A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, ^{131}I, ^{112}In, ^{99}mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of ^{99}mTc. The labeled antibody or antibody fragment will then preferentially accumulate at he location of cells which contain the specific protein. In

vivo tumor imaging is described in S.W. Burchiel et al., „Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments". (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B.A. Rhodes, eds., Masson Publishing Inc. (1982)).

The marker Trp8a and Trp8b is also useful for prognosis, for monitoring the progression of the tumor and the diagnostic evaluation of the degree of malignancy of a prostate tumor (grading and staging), e.g. by using in situ hybridization: In a primary carcinoma Trp8 is expressed in about 2 to 10% of carcinoma cells, in a rezidive carcinoma in about 10 to 60% of cells and in metastases in about 60 to 90% of cells.

The present invention also relates to a method for diagnosing endometrial cancer (cancer of the uterus) which comprises contacting a target sample suspected to contain the protein Trp8a and/or Trp8b or the Trp8a and/or Trp8b encoding mRNA with a reagent which reacts with Trp8a and/or Trp8b or the encoding mRNA and detecting Trp8a and/or Trp8b encoding mRNA. As regards particular embodiments of this method reference is made to the particular embodiments of the method of diagnosing a prostate cancer outlined above.

For evaluating whether the concentration of Trp8a, Trp8b, Trp10a or Trp10b or the concentration of Trp8a, Trp8b, Trp10a or Trp10b encoding mRNA is normal or increased, thus indicative for the presence of a malignant tumor, the measured concentration is compared with the concentration in a normal tissue, preferably by using the ratio of Trp8a:Trp9, Trp8b:Trp9 or Trp10(a or b)/Trp9 for quantification.

Since the prostate carcinoma forms its own basement membrane when growing invasively, it can be concluded that only cells expressing Trp8 and Trp10 are involved in this phenomenon. Thus, it can be concluded that by inhibiting the expression and/or activity of these proteins an effective therapy of cancers like PCA is provided.

Thus, the present invention also relates to a pharmaceutical composition containing a reagent which decreases or inhibits Trp8a, Trp8b, Trp10a and/or Trp10b expression or the activity of Trp8a, Trp8b, Trp10a and/or Trp10b, and a method for preventing, treating, or ameliorating a prostate tumor, endometrial cancer (uterine carcinoma) tumor, a chorion carcinoma, cancer of the lung or melanoma,which comprises administering to a mammalian subject a

therapeutically effective amount of a reagent which decreases or inhibits Trp8a, Trp8b, Trp10a and/or Trp10b expression or the activity of Trp8a, Trp8b, Trp10a and/or Trp10b. Examples of such reagents are the above described antisense RNAs, ribozymes or inhibitors, e.g. specific antibodies. Furthermore, peptides, which inhibit or modulate the biological function of Trp8a, Trp8b, Trp9, Trp10a and/or Trp10b may be useful as therapeutical reagents. For example, these peptides can be obtained by screening combina torial phage display libraries (Cosmix, Braunschweig, Germany) as described by Rottgen, P. and Collins, J. (Gene (1995) 164 (2): 243-250). Furthermore, antigenic epitopes of the Trp8 and Trp10 proteins can be identified by the expression of recombinant Trp8 and Trp10 epitope libraries in E. coli (Marquart, A. & Flockerzi, V., FEBS Lett. 407 (1997), 137-140; Trost, C., et al., FEBS Lett. 451 (1999) 257-263 and the consecutive screening of these libraries with serum of patients with cancer of the prostate or of the endometrium. Those Trp8 and Trp10 epitopes which are immunogenic and which lead to the formation of antibodies in the serum of the patients can be then be used as Trp8 or Trp10 derived peptide vaccines for immune inventions against cancer cells which express Trp8 or Trp10. Alternatively to the E. coli expression system, Trp8 or Trp10 or epitopes of Trp8 and Trp10 can be expressed in mammalian cell lines such as human embryonic kidney (Hek 293) cells (American Type Culture Collection, ATCC CRL 1573).

Finally, compounds useful for therapy of the above described diseases comprise compounds which act as antagonists or agonists on the ion channels Trp8, Trp9 and Trp10. It could be shown that Trp8 is a highly calcium selective ion channel which in the presence of monovalent (namely sodium) and divalent ions (namely calcium) is only permeable for calcium ions (see Example 4, below, and Figures 3A, C, E). Under physiological conditions, Trp8 is a calcium selective channel exhibiting large inward currents. This very large conductance of Trp8 channels (as wells as Trp9 and Trp10a/b channels) is useful to establish systems for screening pharmacological compounds interacting with Trp-channels including high throughput screening systems. Useful high throughput screening systems are well known to the person skilled in the art and include, e.g., the use of cell lines stably or transiently transfected with DNA sequences encoding Trp8, Trp9 and Trp10 channels in assays to detect calcium signaling in biological systems. Such systems include assays based on Ca-sensitive dyes such as aequorin, apoaequorin, Fura-2, Fluo-3 and Indo-1.

Accordingly, the present invention also relates to a method for identifying compounds which act as agonists or antagonists on the ion channels Trp8, Trp9 and/or Trp10, said method comprising contacting a test compound with the ion channel Trp8, Trp9 and/or Trp10, preferably by using a system based on cells stably or transiently transfected with DNA sequences encoding Trp8, Trp9 and/or Trp10, and determining whether said test compound affects the calcium uptake.

For administration the above described reagents are preferably combined with suitable pharmaceutical carriers. Examples of suitable pharmaceutical carriers are well known in the art and include phosphate buffered saline solutions, water, emulsions, such as oil/water emulsions, various types of wetting agents, sterile solutions etc. Such carriers can be formulated by conventional methods and can be administered to the subject at a suitable dose. Administration of the suitable compositions may be effected by different ways, e.g. by intravenous, intraperetoneal, subcutaneous, intramuscular, topical or intradermal administration. The route of administration, of course, depends on the nature of the tumor and the kind of compound contained in the pharmaceutical composition. The dosage regimen will be determined by the attending physician and other clinical factors. As is well known in the medical arts, dosages for any one patient depends on many factors, including the patient's size, body surface area, age, sex, the particular compound to be administered, time and route of administration, the kind and stage of the tumor, general health and other drugs being administered concurrently.

The delivery of the antisense RNAs or ribozymes of the invention can be achieved by direct application or, preferably, by using a recombinant expression vector such as a chimeric virus containing these compounds or a colloidal dispersion system. By delivering these nucleic acids to the desired target, the intracellular expression of Trp8a, Trp8b, Trp10a and/or Trp10b and, thus, the level of Trp8a, Trp8b, Trp10a and/or Trp10b can be decreased resulting in the inhibition of the negative effects of Trp8a, Trp8b, Trp10a and/or Trp10b, e.g. as regards the metastasis formation of PCA.

Direct application to the target site can be performed, e.g., by ballistic delivery, as a colloidal dispersion system or by catheter to a site in artery. The colloidal dispersion systems which can be used for delivery of the above nucleic, acids include macromolecule complexes, nanocapsules, microspheres, beads and lipid-based systems including oil-in-water emulsions

(mixed), micelles, liposomes and lipoplexes, The preferred colloidal system is a liposome. The composition of the liposome is usually a combination of phospholipids and steroids, especially cholesterol. The skilled person is in a position to select such liposomes which are suitable for the delivery of the desired nucleic acid molecule. Organ-specific or cell-specific liposomes can be used in order to achieve delivery only to the desired tumor. The targeting of liposomes can be carried out by the person skilled in the art by applying commonly known methods. This targeting includes passive targeting (utilizing the natural tendency of the liposomes to distribute to cells of the RES in organs which contain sinusoidal capillaries) or active targeting (for example by coupling the liposome to a specific ligand, e.g., an antibody, a receptor, sugar, glycolipid, protein etc., by well known methods). In the present invention monoclonal antibodies are preferably used to target liposomes to specific tumors via specific cell-surface ligands.

Preferred recombinant vectors useful for gene therapy are viral vectors, e.g. adenovirus, herpes virus, vaccinia, or, more preferably, an RNA virus such as a Retrovirus. Even more preferably, the retroviral vector is a derivative of a murine or avian retrovirus. Examples of such retroviral vectors which can be used in the present invention are: Moloney murine leukemia virus (MoMuLV), Harvey murine sarcoma virus (HaMuSV), murine mammary tumor virus (MuMTV) and Rous sarcoma virus (RSV). Most preferably, a non-human primate retroviral vector is employed, such as the gibbon ape leukemia virus (GaLV), providing a broader host range compared to murine vectors. Since recombinant retroviruses are defective, assistance is required in order to produce infectious particles. Such assistance can be provided, e.g., by using helper cell lines that contain plasmids encoding all of the structural genes of the retrovirus under the control of regulatory sequences within the LTR. Suitable helper cell lines are well known to those skilled in the art. Said vectors can additionally contain a gene encoding a selectable marker so that the transduced cells can be identified. Moreover, the retroviral vectors can be modified in such a way that they become target specific. This can be achieved, e.g., by inserting a polynucleotide encoding a sugar, a glycolipid, or a protein, preferably an antibody. Those skilled in the art know additional methods for generating target specific vectors. Further suitable vectors and methods for in vitro- or in vivo-gene therapy are described in the literature and are known to the persons skilled in the art; see, e.g., WO 94/29469 or WO 97/00957.

In order to achieve expression only in the target organ, i.e. tumor to be treated, the nucleic acids encoding, e.g. an antisense RNA or ribozyme can also be operably linked to a tissue specific promoter and used for gene therapy. Such promoters are well known to those skilled in the art (see e.g. Zimmermann et al., (1994) Neuron 12, 11-24; Vidal et al.; (1990) EMBO J. 9, 833-840; Mayford et al., (1995), Cell 81, 891-904; Pinkert et al., (1987) Genes & Dev. 1, 268-76).

For use in the diagnostic research discussed above, kits are also provided by the present invention. Such kits are useful for the detection of a target cellular component, which is Trp8a, Trp8b, Trp10a and/or Trp10b or, alternatively, Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA or Trp10a/b antisense transcripts, wherein the presence or an increased concentration of Trp8a, Trp8b, Trp10a and/or Trp10b or, alternatively, Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA or Trp10a/b antisense transcripts is indicative for a prostate tumor, endometrial cancer, melanoma, chorion carcinoma or cancer of the lung, said kit comprising a probe for detection of Trp8a, Trp8b, Trp9, Trp10a and/or Trp10b or, alternatively, Trp8a, Trp8b, Trp9, Trp10a and/or Trp10b encoding mRNA or Trp10a/b antisense transcripts. The probe can be detectably labeled. Such probe may be a specific antibody or specific oligonucleotide. In a preferred embodiment, said kit contains an anti-Trp8a-, anti-Trp8b-, anti-Trp9-, anti-Trp10a-and/or anti-Trp10b-antibody and allows said diagnosis, e.g., by ELISA and contains the antibody bound to a solid support, for example, a polystyrene microtiter dish or nitrocellulose paper, using techniques known in the art. Alternatively, said kits are based on a RIA and contain said antibody marked with a radioactive isotope. In a preferred embodiment of the kit of the invention the antibody is labeled with enzymes, fluorescent compounds, luminescent compounds, ferromagnetic probes or radioactive compounds. The kit of the invention may comprise one or more containers filled with, for example, one or more probes of the invention. Associated with container (s) of the kit can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, us or sale for human administration.

EXAMPLES

The following Examples are intended to illustrate, but not to limit the invention. While such Examples are typical of those that might be used, other methods known to those skilled in the art may alternatively be utilized.

Example 1: Materials and Methods

(A) Isolation of cDNA clones and Northern blot analysis

Total RNA was isolated from human placenta an prostate using standard techniques. Isolation of mRNA was performed with poly $(A)^+$RNA - spin columns (New England Biolabs, Beverly, USA) according to the instructions of the manufacturer. Poly (a) $^+$RNA was reverse transcribed using the cDNA choice system (Gibco-BRL, Rockville, USA) and subcloned in λ-Zap phages (Stratagene, La Jolla, USA). An human expressed sequence tag (GenBank accession number 1404042) was used to screen an oligo d(T) primed human placenta cDNA library. Several cDNA clones were identified and isolated. Additional cDNA clones were isolated from two specifically primed cDNA libraries using primers 5'-gca tag gaa ggg aca ggt gg-3' and 5'-gag agt cga ggt cag tgg tcc-3'.

cDNA clones were sequenced using a thermocycler (PE Applied Biosystems, USA) and Thermo Sequenase (Amersham Pharmacia Biotech Europe, Freiburg, Germany). DNA sequences were analyzed with an automated sequencer (Licor, Linccoln, USA).

For Northern blot analysis 5 µg human poly $(A)^+$ RNA from human placenta or prostate were separated by electrophoresis on 0.8 % agarose gels. Poly $(A)^+$ RNA was transferred to Hybond N nylon membranes (Amersham Pharmacia Biotech Europe, Freiburg, Germany). The membranes were hyridized in the presence of 50 % formamide at 42^0C over night. DNA probes were labelled using [α^{32}P]dCTP and the „ready prime„ labelling kit (Amersham Pharmacia Biotech Europe, Freiburg, Germany). Commercial Northern blots were hybridized according to the distributors instructions (Clontech, Paolo Alto, USA).

(B) Construction of expression plasmids and transfection of HEK 293 cells

Lipofections were carried out with the recombinant dicistronic eucaryotic expression plasmid pdiTRP8 containing the cDNA of Trp8b under the control of the chicken ß-actin promotor followed by an internal ribosome entry side (IRES) and the cDNA of the green fluorescent protein (GFP). To obtain pdiTRP8 carrying the entire protein coding regions of TRP8b and

the GFP (Prasher, D.C. et al. (1992), Gene 111, 229-233), the 5`and 3`-untranslated sequences of the TRP8b cDNA were removed, the consensus sequence for initiation of translation in vertebrates (Kozak, M. (1987) Nucleic Acids Research15, 8125-8148) was introduced immediately 5`of the translation initiation codon and the resulting cDNA was subcloned into the pCAGGS vector (Niwa, H., Yamamura, K. and Miyazaki, J (1991), Gene 8, 193-199) downstream of the chicken β-actin promotor. The IRES derived from encephalmyocarditis virus (Kim, D.G., Kang, H.M., Jang, S.K. and Shin H.S. (1992) Mol.Cell.Biol. 12, 3636-3643) followed by the GFP cDNA containing a Ser65Thr mutation (Heim, R., Cubitt, A.B., Tsien, R.Y. (1995) Nature 373, 663-664) was then cloned 3` to the TRP8b cDNA. The IRES sequence allows the simultaneous translation of TRP8b and GFP from one transcript. Thus, transfected cells can be detected unequivocally by the development of green fluorescence.

For monitoring of the intracellular Ca^{2+} concentration human embryonic kidney (HEK 293) cells were cotransfected with the pcDNA3-TRP8b vector and the pcDNA3-GFPvector in a molar ratio of 4 : 1 in the presence of lipofectamine (Quiagen, Hilden, Germany). To obtain pcDNA3-TRP8b the entire protein coding region of TRP8b including the consensus sequence for initiation of translation in vertebrates (Kozak, M. (1987) Nucleic Acids Research15, 8125-8148) was subcloned into the pcDNA3 vector (Invitrogen, Groningen, Netherlands). Calcium monitoring and patch clamp experiments were carried out two days and one day after transfection, respectively.

(C) Chromosomal localization of the Trp8 gene

The chromosomal localization of the human TRP8 gene was performed using NIGMS somatic hybrid mapping panel No.2 (Coriell Institute, Camden, NJ, USA) previously described (Drwinga, H.L., Toji, L.H., Kim, C.H., Greene, A.E., Mulivor, R.A. (1993) Genomics 16, 311-314; Dubois, B.L. and Naylor, S.L. (1993) Genomics 16, 315-319).

(D) In Vitro Translation, glutathione - sepharose and calmodulin agarose binding assay

N– and C-terminal Trp8-fragments were subcloned into the pGEX-4T2 vector (Amersham Pharmacia Europe, Freiburg, Germany) resulting in glutathione-S-transferase (GST)-Trp8 fusion constructs (Fig. 4). The GST-TRP8-fusion proteins were expressed in E. coli BL 21 cells and purified using glutathione - sepharose beads (Amersham Pharmacia Biotech Europe, Freiburg, Germany).

In vitro translation of human Trp8 cDNA and Xenopus laevis calmodulin cDNA (Davis, T.N. and Thorner, J. Proc.Natl.Acad.Sci. USA 86, 7909-7913.) was performed in the presence of ^{35}S-methionine using the TNT coupled transcription/translation kit (Promega, Madison, USA). Translation products were purified by gel fliltration (Sephadex G50, Amersham Pharmacia Biotech Europe, Freiburg, Germany) and equal amounts of ^{35}S labeled probes were incubated for 2 h with glutathione beads bound to GST - Trp8 or calmodulin - agarose (Calbiochem) in 50 mM Tris-HCl, pH 7.4, 0.1 % Triton X-100, 150 mM NaCl in the presence of 1 mM Ca^{2+} or 2 mM EGTA. After three washes, bound proteins were eluted with SDS sample buffer, fractionated by SDS-PAGE and ^{35}S labeled proteins were detected using a Phosphor Imager (Fujifilm, Tokyo, Japan).

(E) Calcium measurements

The intracellular Ca^{2+} concentration ($[Ca^{2+}]_i$) was determined by dual wavelength fura-2 fluorescence ratio measurements (Tsien, R.Y. (1988) Trends Neurosci. 11, 419-424) using a digital imaging system (T.I.L.L. Photonics, Planegg, Germany). HEK cells were grown in minimal essential medium in the presence of 10 % fetal calf serum and cotransfected with the pcDNA3-TRP8b vector and the pCDNA3-GFPvector as described above (B). Transfected cells were detected by development of green fluorescence. The cells were loaded with 4μM fura-2/AM (Molecular Probes, Oregon, USA) for one hour. After loading the cells were rinsed 3 times with buffer B1 (10 mM Hepes, 115 mM NaCl, 2 mM $MgCl_2$, 5mM KCl, pH 7.4) and the $[Ca^{2+}]_i$ was calculated from the fluorescence ratios obtained at 340 and 380 nm excitation wavelengths as described (Garcia, D.E., Cavalié, A. and Lux, H.D. (1994) J. Neurosci 14, 545-553).

(F) Electrophysiological recordings

HEK cells were transfected with the eucaryotic expression plasmid pdiTRP8 described in (B) and electrophysiolocigal recordings were carried out one day after transfection. Single cells were voltage clamped in the whole cell mode of the patch clamp technique as described (Hamill, O.P., Marty, A., Neher, E., Sakmann, B. and Sigworth, F.J. (1981) Pflügers Arch. 391, 85-100; Philipp, S., Cavalié, A., Freichel, M., Wissenbach, U., Zimmer, S., Trost, C., Marquart, A., Murakami, M. and Flockerzi, V. (1996) EMBO J. 6166-6171). The pipette solution contained contained (mM): 140 aspartic acid, 10 EGTA, 10 NaCl, 1 MgCl2, 10 Hepes (pH 7.2 with CsOH) or 125 CsCl, 10 EGTA, 4 $CaCl_2$ 10 Hepes (pH 7,2 with CsOH). The bath solution contained (mM): 100 NaCl, 10 CsCl, 2 $MgCl_2$, 50 mannitol, 10 glucose, 20

Hepes (pH 7,4 with CsOH) and 2 $CaCl_2$, or no added $CaCl_2$ (-Ca^{2+} solution). Divalent free bath solution contained (mM): 110 N-methyl-D-glucamine (NMDG). Whole cell currents were recorded during 100 msec voltage ramps from –100 to +100 mV at varying holding potentials.

(G) In Situ Hybridization

In situ hybridizations were carried out using formalin fixed tissue slices of 6 - 8 µM thickness. The slices were hydrated and incubated in the presence of PBS buffer including 10 µg / ml proteinase K (Roche Diagnostics, Mannheim, Germany) for 0.5 h. The slices were hybridized at 37°C using biotinylated deoxy-oligonucleotides (0.5 pmol / µl) in the presence of 33 % formamide for 12 h. Furthermore the slices were several times rinsed with 2 x SSC and incubated at 25°C for 0.5 h with avidin / biotinylated horse raddish peroxidase complex (ABC, DAKO, Santa Barbara, USA). After several washes with PBS buffer the slices were incubated in the presence of biotinylated tyramid and peroxide (0.15 % w/v) for 10 min, rinsed with PBS buffer and additionally incubated with ABC complex for 0.5 h. The slices were washed with PBS buffer and incubated in the presence of DAB solution (diaminobenzidine (50µg / ml), 50 mM Tris/EDTA buffer pH 8.4, 0.15 % H_2O_2 in N,N - dimethyl-formamide; Merck, Darmstadt, Germany), The detection was stopped after 4 minutes by incubating the slides in water. Tyramid was biotinylated by incubating NHS-LC Biotin (sulfosuccinimidyl-6-(biotinimid)-hexanoat), 2.5 mg / ml; Pierce, Rockford, USA) and tyramin-HCl (0.75 mg / ml, Sigma) in 25 mM borate buffer pH 8.5 for 12 h. The tyramid solution was diluted 1 - 5 : 1000 in PBS buffer.

(H) GenBank accession numbers: TRP8a, Aj243500; TRP8b Aj243501

Example 2: Expression of TRP8 transcripts

In search of proteins distantly related to the TRP family of ion channels, an human expressed sequence tag (EST, GenBank accession number 1404042) was identified in the GenBank database using BLAST programms (at the National Center for Biotechnology Information (NCBI); Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J.J. (1990) Mol. Biol. 5, 403-410) being slightly homologous to the VR1 gene. Several human placenta cDNA libraries were constructed and screeened with this EST DNA as probe. Several full length

cDNA clones were identified and isolated. The full length cDNA clones encoded two putative proteins differing in three amino acids and were termed Trp8a and Trp8b (Fig. 1c, 2a, 7 and 8A). This finding was reproduced by isolating cDNA clones from two cDNA libraries constructed from two individual placentas. The derived protein sequence(s) comprises six transmembrane domains, a characteristic overall feature of trp channels and related proteins (Fig.: 1b). The sequence is closely related to the meanwhile published calcium uptake transport protein 1 (CaT1), isolated from rat intestine (Peng, J.B., Chen, X.Z., Berger, U.V., Vassilev, P.M., Tsukaguchi, H., Brown, E.M. and Hediger M.A.(1999) J Biol Chem. 6;274, 22739-22746) and to the epithelial calcium uptake channel (ECaC) isolated from rabbit kidney (Hoenderop, J.G., van der Kemp, A.W., Hartog, A., van de Graaf, S.F., van Os, C.H., Willems, P.H. and Bindels, R.J. (1999) J Biol Chem. 26;274, 8375-8378). Expression of Trp8a/b transcripts are detectable in human placenta, pancreas and prostate (Fig.: 5) and the size of the Northern signal (3.0 kb) corresponds with the size of the isolated full length cDNAs. In addition, a shorter transcript of 1.8 kb, probably a splice variant, is detectable in human testis. The Trp8 mRNA is not expressed in small intestine or colon (Fig.: 5) implicating that Trp8 is not the human ortholog of the rat CaT1 or rabbit ECaC proteins. To investigate whether there are other related sequences Trp8a/b derived primers (UW241, 5'-TAT GAG GGT TCA GAC TGC-3' and UW242, 5'-CAA AGT AGA TGA GGT TGC-3') were used to amplify a 105 bp fragment from human genomic DNA being 95% identical on the nucleotide level to the Trp8 sequence (data not shown). This indicates the existence of several similar sequences in humans at least at the genomic level.

Example 3: Two variants of the Trp8 protein (Trp8a and Trp8b) arise by polymorphism

Two variants of the Trp8 cDNA were isolated from human placenta (Fig.: 2A, 7 and 8A) which encoded two proteins which differ in three amino acids and were termed Trp8a and Trp8b. Trp8a/b specific primers were designed to amplify a DNA fragment of 458 bp of the Trp8 gene from genomic DNA isolated from human T-lymphocytes (primer pair: UW243, 5'-CAC CAT GTG CTG CAT CTA CC-3' and UW244, 5'-CAA TGA CAG TCA CCA GCT CC-3'). The amplification product contains a part of the sequence where the derived protein sequence of Trp8a comprises the amino acid valine and the Trp8b sequence methionine as well as a silent base pair exchange (g versus a) and an intron of 303bp (Fig.: 2.A, B). Both variants of the Trp8 genes (a,b) were amplified from genomic DNA in equal amounts indicating the existence of both variants in the human genome and therefore being not the

result of RNA editing (Fig.: 2B). The Trp8a gene can be distinguished from the Trp8b gene by cutting the genomic fragment of 458bp with the restriction enzyme Bsp1286I (Fig. 2B). Using human genomic DNA isolated from blood of twelve human subjects as template, the 458bp fragment was amplified and restricted with BSP1286I. In eleven of the tested subjects only the Trp8b gene is detectable, while one subject (7) contains Trp8a and Trp8b genes (Fig.: 2D). These implicates that the two Trp8 variants arise by polymorphism and do not represent individual genes. Using Trp8 specific primers and chromosomal DNA as template, the Trp8 locus is detectable on chromosome 7 (Fig.: 2C).

Example 4: Trp8b is a calcium permeable channel

The protein coding sequence of the Trp8b cDNA was subcloned into pcDNA3 vector (Invitrogen, Groningen, Netherlands) under the control of the cytomegalovirus promotor (CMV). Human embryonic kidney (HEK 293) cells were cotransfected with the Trp8b pcDNA3 construct (pcDNA3-Trp8b vector) and the pcDNA3-GFPvector encoding the green fluorescent protein (GFP) in 4:1 ratio. The Trp8b cDNA and the cDNA of the reporter, GFP, was transiently expressed in human embryonic kidney (HEK 293) cells. The intracellular Ca^{2+} concentration ($[Ca^{2+}]_i$) and changes of $[Ca^{2+}]_i$ were determined by dual wavelength fura-2 fluorescence ratio measurements (Fig.: 3F) in cotransfected cells which were identified by the green fluorescence of the reporter gene GFP.

Dual wavelength fura-2 fluorescence ratio measurement is a standard procedure (e.g. in: An introduction of Molecular Neurobiology (ed. Hall, Z.W.)Sinauer Associates, Sunderland, USA (1992)) using fura-2, which is a fluorescent Ca^{2+} sensitive dye and which was designed by R.Y.Tsien (e.g. Trends Neurosci. 11, 419-424 (1988) based upon the structure of EGTA. Its fluorescence emission spectrum is altered by binding to Ca^{2+} in the physiological concentration range. In the absence of Ca^{2+}, fura-2 fluoresces most strongly at an excitation wavelength of 385 nm; when it binds Ca^{2+}, the most effective excitation wavelength shifts to 345 nm. This property is used to measure local Ca^{2+} concentrations within cells. Cells can be loaded with fura-2 esters (e.g. fura-2AM) that diffuse across cell membranes and are hydrolyzed to active fura-2 by cytosolic esterases.

In the presence of 1mM Ca^{2+}, Trp8 expressing cells typically contained more than 300 nM cytosolic Ca^{2+}, while non transfected controls contained less than 100 nM Ca^{2+} ions (Fig. 3F).

When Trp8b transfected cells were incubated without extracellular Ca^{2+}, the intracellular Ca^{2+} concentration ($[Ca^{2+}]_i$) decreased to levels comparable to non transfected cells. Readdition of 1mM Ca^{2+} to the bath resulted in significant increase of the cytosolic $[Ca^{2+}]$ in Trp8b transfected cells, but not in controls (Fig.: 3F). After readdition of Ca^{2+} ions to the bath solution, the cytosolic Ca^{2+} concentration remains on a high steady state level in the Trp8b transfected cells.

Example 5: Trp8 expressing cells show calcium selective inward currents

To characterize in detail the electrophysiological properties of TRP8, TRP8 and GFP were coexpressed in HEK293 cells using the dicistronic expression vector pdiTRP8 and measured currents using the patch clamp technique in the whole cell mode (Hamill, O.P., Marty, A., Neher, E., Sakmann, B. and Sigworth, F.J. (1981) Pflugers Arch., 391, 85-100).

The eucaryotic expression plasmid pdiTRP8 contains the cDNA of Trp8b under the control of the chicken ß-actin promotor followed by an internal ribosome entry side (IRES) and the cDNA of the green fluorescent protein (GFP). To obtain pdiTRP8 carrying the entire protein coding regions of TRP8b and the GFP (Prasher, D.C. et al. (1992), Gene 111, 229-233), the 5`and 3´-untranslated sequences of the TRP8b cDNA were removed, the consensus sequence for initiation of translation in vertebrates (Kozak, M. (1987) Nucleic Acids Research15, 8125-8148) was introduced immediately 5`of the translation initiation codon and the resulting cDNA was subcloned into the pCAGGS vector (Niwa, H., Yamamura, K. and Miyazaki, J (1991), Gene 8, 193-199) downstream of the chicken β-actin promotor. The IRES derived from encephalmyocarditis virus (Kim, D.G., Kang, H.M., Jang, S.K. and Shin H.S. (1992) Mol.Cell.Biol. 12, 3636-3643) followed by the GFP cDNA containing a Ser65Thr mutation (Heim, R., Cubitt, A.B., Tsien, R.Y. (1995) Nature 373, 663-664) was then cloned 3` to the TRP8b cDNA. The IRES sequence allows the simultaneous translation of TRP8b and GFP from one transcript. Thus, transfected cells can be detected unequivocally by the development of green fluorescence.

In the presence of 2 mM external calcium, Trp8b transfected HEK cells show inwardly rectifying currents, the size of which depends on the level of intracellular calcium and the electrochemical driving force. The resting membrane potential was held either at –40 mV, or, to lower the driving force for calcium influx in between pulses, at + 70 mV. Current traces

were recorded in response to voltage ramps from –100 to +100 mV, that were applied every second. To monitor inward and outward currents over time, we analyzed the current size at – 80 and + 80 mV of the ramps. Figure 3A shows a representative trace of the current at – 80 mV over time. Both at a holding potential of –40 mV or at +70 mV, the currents are significantly larger than in cells transfected with only the GFP containing vector (Fig.: 3E). Interestingly, after changing to a positive holding potential, current size in Trp8 transfected cells slowly increases and reaches steady state after approximately 70 seconds (Fig.: 3A). To determine the selectivity of the induced currents, we then perfused the cells with solutions that either contain no sodium, no added Ca^{2+} (Fig. 3A, C) or a sodium containing, but divalent ion free bath solution. To control for the effect of the solution change alone, we also perfused with normal bath (see puff in Fig. 3A). While removal of external Ca^{2+} completely abolishes the trp 8 induced currents - the remaining current being identical in size and shape to the control (Fig.: 3A, C, E), removal of external sodium has no effect (Fig.: 3E). An important hallmark of calcium selective channels (e.g. Vennekens, R., Hoenderop, G.J., Prenen, J., Stuiover, M., Willems, PHGM, Droogmans, G., Nilius, B.and Bindels, R.J.M (1999) J. Biol. Chem. 275, 3963-3969), is their ability to conduct sodium only if all external divalent ions, namely Ca^{2+} and magnesium are removed. To test whether the trp 8 channel conforms with this phenomenon normal bath solution was switched to a solution containing only sodium and 1 mM EGTA. As can be seen in Figure 3B and D, Trp8 transfected cells can now conduct very large sodium currents. Interestingly, immediately after the solution change, the currents first become smaller before increasing rapidly, indicating that the pore may initially still be blocked by calcium a phenomenon usually called anomalous mole fraction behaviour (Warnat, J., Philipp, S., Zimmer, S.,.Flockerzi, V., and Cavalié A.(1999) J.Physiol. (Lond) 518, 631-638). The measured outward currents of Trp8 transfected cells in normal bath solution are not significantly different from non-transfected control cells or cells which only express the reporter gene GFP. As the removal of external Ca^{2+} abolishes the Trp8 specific current, the remaining current was subtracted from the current before the solution change to obtain the uncontaminated Trp8 conductance (see inset in Fig.: 3C). As expected from the given ionic conditions (high EGTA inside, 2 mM Ca^{2+} outside), the current-voltage relationship now shows prominent inward rectification with little to no outward current.

Both the time course of the development of Trp8 currents and the size of the currents depend on the frequency of stimulation (data not shown), the internal and external Ca^{2+} concentration

and the resting membrane potential, suggesting that Trp8 calcium conductance is intrically regulated by a Ca^{2+} mediated feedback mechanisms.

Example 6: Ca^{2+} / calmodulin binds to the C-terminus of the Trp8 protein

To test whether calmodulin, a prime mediator of calcium regulated feedback, is involved, first it was investigated biochemically whether Trp8 protein can bind calmodulin. Trp8 cDNA was in vitro translated in the presence of ^{35}S-methionine and the product incubated with calmodulin-agarose beads. After several washes either in the presence or abscence of Ca^{2+}, the beads were incubated in Laemmli buffer and subjected to SDS-polyacrylamide gel electrophoresis. In the presence of Ca^{2+} (1mM), but not in the absence of Ca^{2+}, Trp8 protein binds to calmodulin (Fig.: 4B).

To narrow down the binding site, two approaches were undertaken: Firstly, GST-TRP8 fusion proteins of various intracellular domains of Trp8 were constructed, expressed in E. coli and bound to gluthathione sepharose beads. These beads were then incubated with in vitro translated ^{35}S- labeled calmodulin, washed and subjected to gel electrophoresis. Secondly, truncated versions of in vitro translated Trp8 protein were used in the above described binding to calmodulin-agarose. As shown in Figure 4A, and C, fusion proteins of the N-terminal region (N1, N2) of Trp8 did not bind calmodulin, while C-terminal fragments (C1, C2, C3, C4) showed calmodulin binding in the presence of calcium (for localization of fragments within the entire Trp8 protein see Fig. 4C). Accordingly, a truncated version of in vitro translated Trp8, which lacks the C-terminal 32 amino acid residues did not bind to calmodulin-agarose (4B). We have restricted the calmodulin binding site to amino acid residues 691 to 711 of the Trp8 protein. This calmodulin binding site does not resemble the typical conserved IQ - motif of conventional myosins, but has limited sequence homology to the calcium dependent calmodulin binding site 1 of the transient receptor potential like (trpl) protein of Drosophila melanogaster (Warr and Kelly, 1996) with several charged amino acid residues conserved. The sequence of the calmodulin binding site of the Trp8 protein resembles a putative amphipathic α-helical wheel structure with a charged and a hydrophobic site according to a model proposed by Erickson-Vitanen and De Grado (1987, Methods Enzymol. 139, 455-478.).

Example 7: Expression of Trp8 transcripts in human placenta and pancreas

Several slides from a human placenta of a ten week old abort were used for in situ hybridization experiments. The in situ hybridization experiments revealed expression of Trp8 transcripts in human placenta (Fig.: 5B). Expression was detectable in trophoblasts and syncytiotrophoblasts of the placenta, but not in Langhans cells.

Trp8 transcripts are detectable in human pancreas (Fig.: 5A). Therefore Trp8 probes were hybridized to tissue sections of human pancreas. The pancreatic tissues were removed from patients with pancreas cancer. Trp8 expression is detectable in pancreatic acinar cells, but not in Langerhans islets (Fig.: 5C). No Trp8 expression was found in regions of pancreatic carcinomas (data not shown).

Furthermore, the Trp8 cDNA is not detectable in human colon nor in human kidney by in situ hybridization as well as by Northern analysis (Fig.: 5A, D). The Northern results taken together with the in situ expression data indicate that the Trp8 protein is not the human ortholog of the CaT1 and ECaC channels cloned from rat intestine (Peng, J.B., Chen, X.Z., Berger, U.V., Vassilev, P.M., Tsukaguchi, H., Brown, E.M. and Hediger M.A.(1999) J Biol Chem. 6;274, 22739-22746) and from rabbit kidney (Hoenderop, J.G., van der Kemp, A.W., Hartog, A., van de Graaf, S.F., van Os, C.H., Willems, P.H. and Bindels, R.J. (1999) J Biol Chem. 26;274, 8375-8378), respectively. Trp8 is unlikely to represent the human version of CaT1 as its expression is undetectable in the small intestine and colon tissues where CaT1 is abundantly expressed. If, however, Trp8 is the human version of rat CaT1, a second gene product appears to be required for Ca^{2+} uptake in human small intestine and colon attributed to CaT1 in rat small intestine and colon.

Example 8: Differential expression of Trp8 transcripts in benign and malign tissue of the prostate

The Trp8 transcripts are expressed in human prostate as shown by hybridization of a Trp8 probe to a commercial Northern blot (Clontech, Palo Alto, USA) (Fig.: 5A). Trp8 transcripts were not detectable by Northern blot analysis using pooled mRNA of patients with benign prostatic hyperplasia (BPH) (Fig.: 5A, prostate*). To examine Trp8 expression on the cellular

level, sections of prostate tissues were hybridized using Trp8 specific cDNA probes (Table 3). Expression of Trp8 transcripts is not detectable in normal prostate (n = 3), benign hyperplasia (BPH, n = 15) or prostatic intraepithelial neoplasia (PIN, n = 9) (Fig.: 6A, C, E). Trp8 transcripts were only detectable in prostate carcinoma (PCA), although with different expression levels. Low expression levels were found in primary carcinomas (2 - 10 % of the carcinoma cells, n = 8) (Fig.: 7B) . Much stronger expression was detectable in rezidive carcinoma (10 - 60 %) (Fig.: 7D, n = 6) and metastases of the prostate (60 - 90 %, n = 4) (Fig.: 7F). Thus it has to be concluded that the commercial Northern blot used in Fig.: 5A contains not only normal prostate mRNA as indicated by the distributor. According to the distributors instructions the prostate mRNA used for this Northern blot was collected from 15 human subjects in the range of 14 to 60 years of age. This prostate tissue was not examined by pathologic means. Since Trp8 expression is not detectable in normal or benign prostate, this finding implicates that the mRNA used for this Northern blot was extracted in part from prostatic carcinoma tissue. To summarize, Trp8 expression is only detectable in malign prostate and, thus, the Trp8 cDNA is a marker for prostate carcinoma. The results are summarized in Table 4.

Table 3

Trp8 probes used for in situ hybridization:

Probes (antisense)

1.) 5' TCCGCTGCCGGTTGAGATCTTGCC 3'

2.) 5' CTTGCTCCATAGGCAGAGAATTAG 3'

3.) 5' ATCCTCAGAGCCCCGGGTGTGGAA3'

Controls (sense)

1.) 5' GGCAAGATCTCAACCGGCAGCGGA 3'

2.) 5' CTAATTCTCTGCCTATGGAGCAAG 3'

3.) 5' TTCCACACCCGGGGCTCTGAGGAT 3'

Table 4

Prostate	total	negative	positive
normal	3	3	0
BPH	15	15	0
PIN	9	9	0

carcinoma	18	1	17

(B) Differential expression of Trp8 transcripts in benign and malign tissue of the uterus

Moreover it could be shown that Trp8 is expresed in endometrial cancer (also called cancer of the uterus, to be distinguished from uterine sarcoma or cancer of the cervix) whereas no expression was observed in normal uterus tissue. Thus, Trp8 also is a specific marker for the diagnosis of the above cancer (Fig. 12).

Example 9: Characterization of Trp9

The complete protein coding sequence of Trp9 was determined (Fig. 9). Trp 9 transcripts are predominantly expressed in the human prostate and in human colon. As it could be shown by Northern blot analysis, there is no difference of the expression of TRP9 in benigne prostata hyperplasia (BPH, Fig. 13, upper panel left) or prostate carcinoma (Fig. 13, upper panel right). However, Trp9 is useful as a reference marker for prostata carcinoma, i.e. can be used for quantifying the expression level of Trp8. The ratio of the expression of Trp8:Trp9 in patients and healthy individuals is useful for the development of a quantitative assay.

Example 10: Characterization of Trp10

The complete protein coding sequence of TRP10 (a and b) was determined by biocomputing (Fig. 10 and 11). Using a 235 bp fragment of the Trp10 cDNA as probe in Northern blot analysis TRP10 transcripts could only be detected in mRNA isolated from individuals with prostate cancer (Fig. 13, bottom panel) but not in mRNA isolated from benign tissue of the prostate (prostate BPH) nor in mRNA isolated from heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas. The 235 bp cDNA fragment of the Trp10 cDNA was amplified using the primer pair UW248 5'-ACA GCT GCT GGT CTA TTC C-3' and UW249 5'-TAT

GTG CCT TGG TTT GTA CC-3' and prostate cDNA as template. In summary, Trp10a and Trp10b , like TRP8 are also expressed in malignant prostate tissue. So far, its expression could not be observed in any other tissue examined (see above). Thus, Trp 10a and Trp10b are also useful markers which are specific for malignant prostate tissue.

Furthermore, database searches in public databases of the national center for biological information (NCBI) revealed the existence of several expressed sequence tags (EST clones) being in part identical to the Trp10 sequence. These EST clones were originally isolated from cancer tissues of lung, placenta, prostate and from melanoma. These clones include the clones with the following accession numbers: BE274448, BE408880, BE207083, BE791173, AI671853, BE390627. The results demonstrate that cancer cells of these tissues express Trp10 related transcripts whereas no expression of Trp10 transcripts in the corresponding healthy tissues are detectable (Figure 13). Furthermore, it could be shown that in cancer cells of melanoma and prostate cancer Trp10 transcripts are expressed as shown by in situ hybridizations using 4 antisense probes (Figure 14A – E and 13K-O and Table 2, above). Furthermore, it could clearly be shown that cancer cells of these tissues expressing Trp10 transcripts also express Trp10-antisense transcripts as shown in Figure 14F-J, Figure 14P-R and Figure14T by in situ hybridizations using 4 sense probes (Table 2, above). The in situ hybridization experiments demonstrate that detection of a subset of cancer cells derived from carcinoma of lung, placenta, prostate and melanoma is feasible using antisense as well as sense probes complementary to Trp10 transcripts or complementary to Trp10-antisense transcripts, respectively.

The foregoing is meant to illustrate but not to limit the scope of the invention. The person skilled in the art can readily envision and produce further embodiment, based on the above teachings, without undue experimentation.

What Is claimed Is:

1. An isolated nucleic acid molecule encoding the human prostate carcinoma associated protein Trp8a, Trp8b, Trp9, Trp10a or Trp10b or a protein exhibiting biological properties of Trp8a, Trp8b, Trp9, Trp10a or Trp10b and being selected from the group consisting of

(a) a nucleic acid molecule encoding a protein that comprises the amino acid sequence depicted in Figure 7, 8A, 9, 10 or 11;

(b) a nucleic acid molecule comprising the nucleotide sequence depicted in Figure 7, 8A, 9, 10 or 11;

(c) a nucleic acid molecule included in DSMZ Deposit No. DSM 13579, DSM 13580, DSM 13584 , DSM 13581 or DSM....;

(d) a nucleic acid molecule which hybridizes to a nucleic acid molecule specified in (a) to (c);

(e) a nucleic acid molecule the nucleic acid sequence of which deviates from the nucleic sequences specified in (a) to (d) due to the degeneration of the genetic code; and

(f) a nucleic acid molecule, which represents a fragment, derivative or allelic variation of a nucleic acid sequence specified in (a) to (e).

2. A recombinant vector containing the nucleic acid molecule of claim 1

3. The recombinant vector of claim 2 wherein the nucleic acid molecule is operatively linked to regulatory elements allowing transcription and synthesis of a translatable RNA in prokaryotic and/or eukaryotic host cells.

4. A recombinant host cell which contains the recombinant vector of claim 3.

5. The recombinant host cell of claim 4, which is a mammalian cell, a bacterial cell, an insect cell or a yeast cell.

6. An isolated protein exhibiting biological properties of the human prostate carcinoma associated protein Trp8a, Trp8b, Trp9, Trp10a or Trp10b which is encoded by a nucleic acid molecule of claim 1.

7. A recombinant host cell that expresses the isolated protein of claim 6.

8. A method of making an isolated protein exhibiting biological properties of the human prostate carcinoma associated protein Trp8a, Trp8b, Trp9, Trp10a or Trp10b comprising:
(a) culturing the recombinant host cell of claim 6 under conditions such that said protein is expressed; and
(b) recovering said protein.

9. The protein produced by the method of claim 8.

10. An antisense RNA sequence characterized in that it is complementary to an mRNA transcribed from a nucleic acid molecule of claim 1 or a part thereof and can selectively bind to said mRNA or part thereof, said sequence being capable of inhibiting the synthesis of the protein encoded by said nucleic acid molecule.

11. A ribozyme characterized in that it is complementary to an mRNA transcribed from a nucleic acid molecule of claim 1 or a part thereof and can selectively bind to and cleave said mRNA or part thereof, thus inhibiting the synthesis of the protein encoded by said nucleic acid molecule.

12. An inhibitor characterized in that it can suppress the activity of the protein of claim 6.

13. A method for diagnosing a prostate carcinoma which comprises contacting a target sample suspected to contain the protein Trp8a, Trp8b, Trp10a and/or Trp10b or the Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA with a reagent which reacts with Trp8a, Trp8b, Trp10a and/or Trp10b or the Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA and detecting Trp8a, Trp8b, Trp10a and/or Trp10b or the Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA.

14. The method of claim 13, wherein the reagent is a nucleic acid.

15. The method of claim 13, wherein the reagent is an antibody.

16. The method of claim 13, wherein the reagent is detectably labeled.

17. The method of claim 16, wherein the label is selected from the group consisting of a radioisotope, a bioluminescent compound, a chemiluminescent compound, a fluorescent compound, a metal chelate, or an enzyme.

18. A method for diagnosing an endometrial cancer (carcinoma of the uterus) which comprises contacting a target sample suspected to contain the protein Trp8a and/or Trp8b or the Trp8a and/or Trp8b encoding mRNA with a reagent which reacts with Trp8a and/or Trp8b or the Trp8a and/or Trp8a and/or trp8b encoding mRNA and detecting Trp8a and/or Trp8b or the Trp8a and/or Trp8b encoding mRNA.

19. The method of claim 18, wherein the reagent is a nucleic acid.

20. The method of claim 18, wherein the reagent is an antibody.

21. The method of claim 18, wherein the reagent is detectably labeled.

22. The method of claim 21, wherein the label is selected from the group consisting of a radioisotope, a bioluminescent compound, a chemiluminescent compound, a fluorescent compound, a metal chelate, or an enzyme.

23. A method for diagnosing a melanoma, chorion carcinoma, cancer of the lung and of the prostate in a tissue of a subject, comprising contacting a sample with a reagent which detects Trp10a and/or Trp10b antisense RNA or Trp10a and/or Trp10b related antisense RNA.

24. A method for preventing, treating, or ameliorating a prostate tumor, endometrial cancer (carcinoma of the uterus) tumor, a chorion carcinoma, cancer of the lung or melanoma, which comprises administering to a mammalian subject a therapeutically effective amount of a reagent which decreases or inhibits expression of Trp8a, Trp8b, Trp10a and/or Trp10b and/or the activity of Trp8a, Trp8b, Trp10a and/or Trp10b.

25. The method of claim 24, wherein the reagent is a nucleotide sequence comprising an antisense RNA.

26. The method of claim 24, wherein the reagent is a nucleotide sequence comprising a ribozyme.

27. The method of claim 24, wherein the reagent is an inhibitor of Trp8a, Trp8b, Trp10a and/or Trp10b.

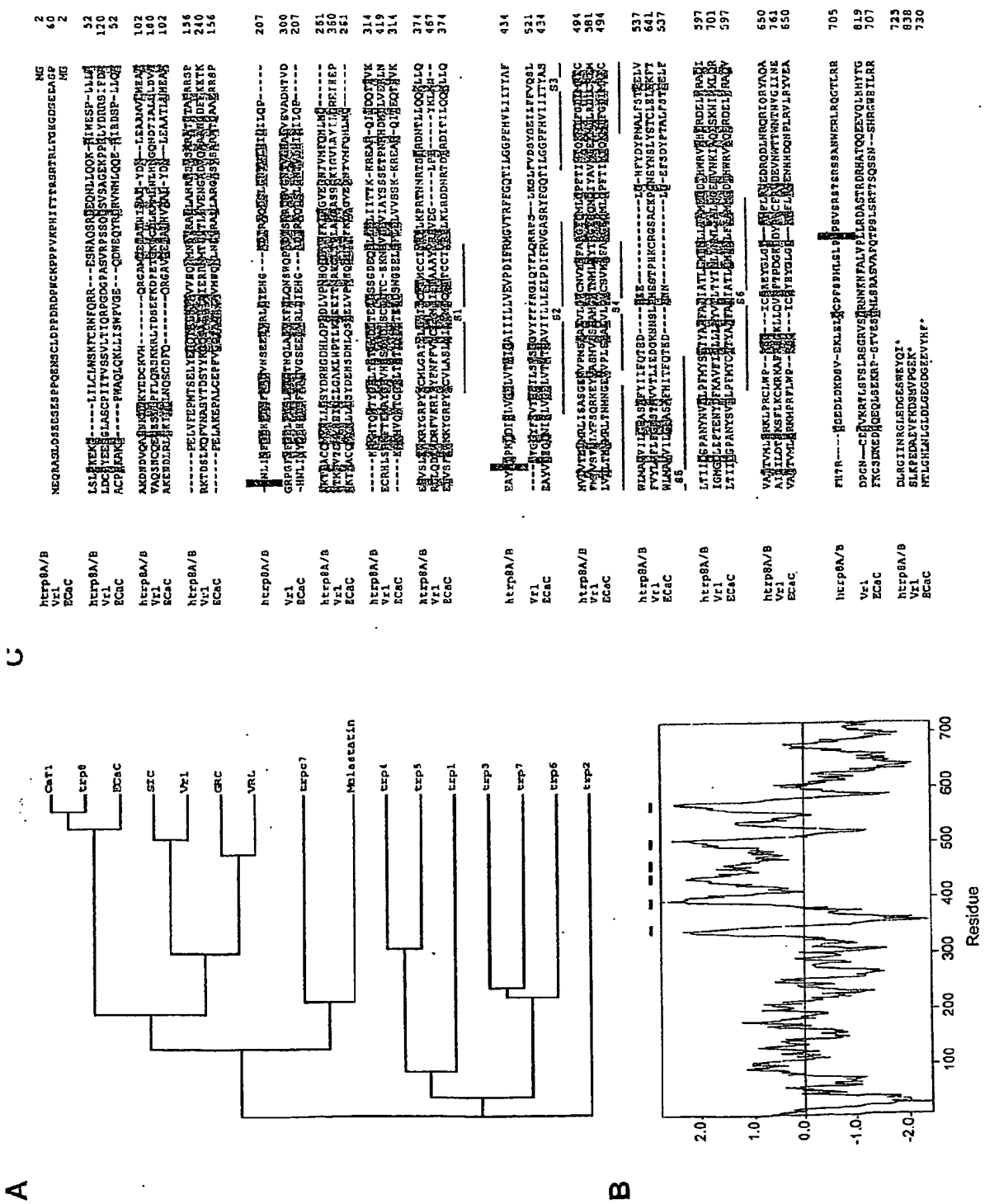
28. The method of claim 27, wherein the reagent is an anti-Trp8a-, anti Trp8b-, anti-Trp10a- and/or anti-Trp10b antibody or a fragment thereof.

29. A diagnostic kit useful for the detection of Trp8a, Trp8b, Trp10a and/or Trp10b or Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA or Trp10a and/or Trp10b antisense transcripts in a sample, wherein the presence of an increased concentration of Trp8a, Trp8b, Trp9, Trp10a and/or Trp10b or Trp8a, Trp8b, Trp9, Trp10a and/or Trp10b encoding mRNA or Trp10a and/or Trp10b antisense transcripts is indicative for a prostate tumor, endometrial cancer (cancer of the uterus) tumor, a chorion carcinoma, cancer of the lung or melanoma, said kit comprising a probe for detection of Trp8a, Trp8b, Trp9, Trp10a or Trp10b or Trp8a, Trp8b, Trp9, Trp10a and/or Trp10b encoding mRNA or Trp10a and/or Trp10b antisense transcripts.

30. The kit of claim 29, wherein the target component to be detected is Trp8a, Trp8b, Trp9, Trp10a and/or Trp10b and the probe is an antibody.

31. A method for identifying a compound which acts as an agonist or antagonist on the ion channels Trp8, Trp9 and/or Trp10, said method comprising contacting a test compound with the ion channel Trp8, Trp9 and/or Trp10,and determining whether said test compound affects the calcium uptake.

Fig. 1


A
Cat1
trp8
BCaC
SrC
Vrl
GRC
VRL
trpc7
Malastatin
trp4
trp5
trp1
trp3
trp7
trp6
trp2
B
2.0
1.0
0.0
-1.0
-2.0
100
200
300
400
500
600
700
Residue
C
htrp8A/B
Vrl
BCaC

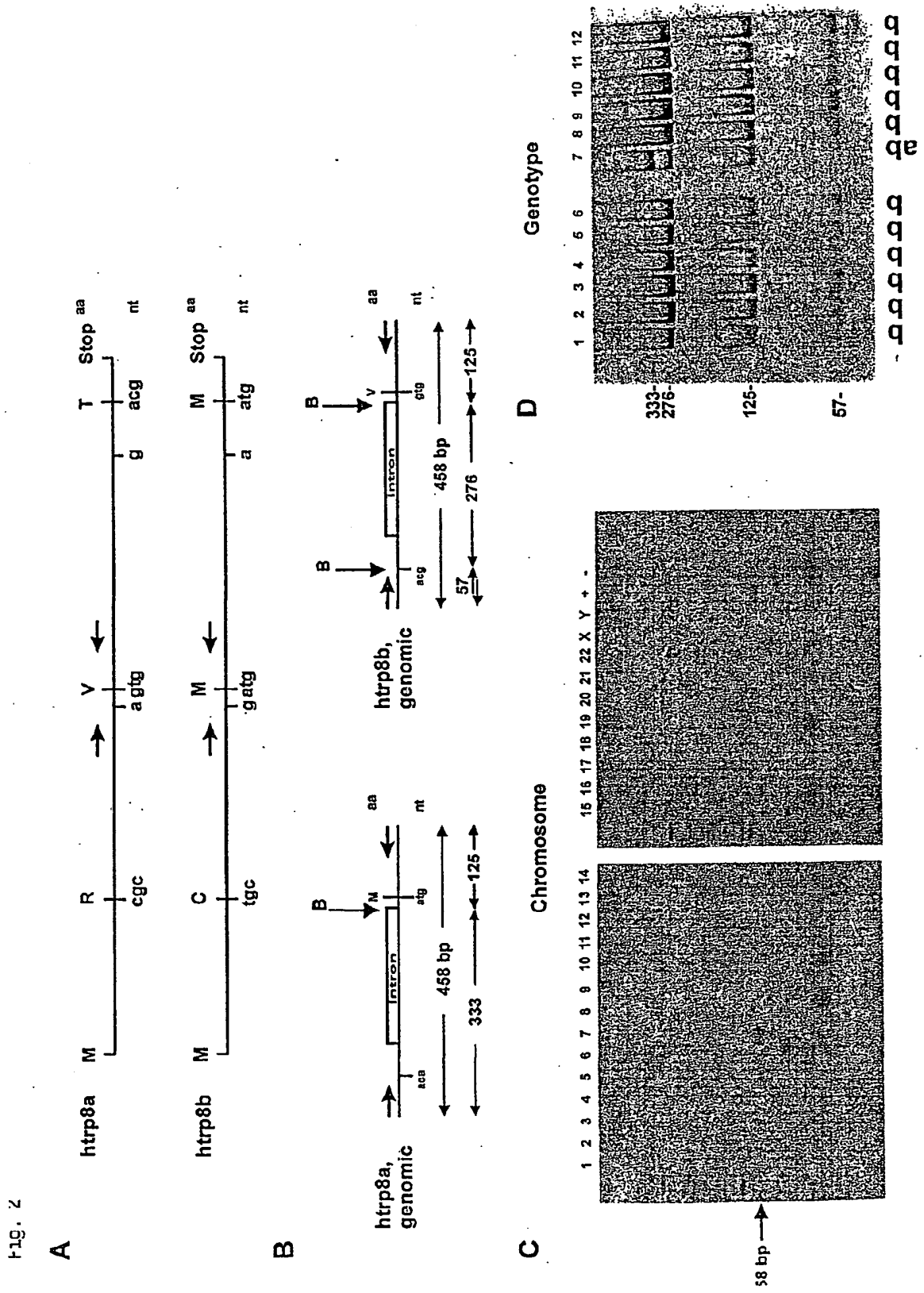

A
htrp8a
M
R
V
T
Stop
aa
nt
cgc
agtg
acg
htrp8b
M
C
M
Stop
tgc
gatg
atg
B
htrp8a, genomic
htrp8b, genomic
Intron
458 bp
333
125
276
57
C
Chromosome
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y + -
58 bp
D
Genotype
1 2 3 4 5 6 6 7 8 9 10 11 12
333-
276-
125-
57-
aa
ab
bb

Fig. 2

Fig. 3

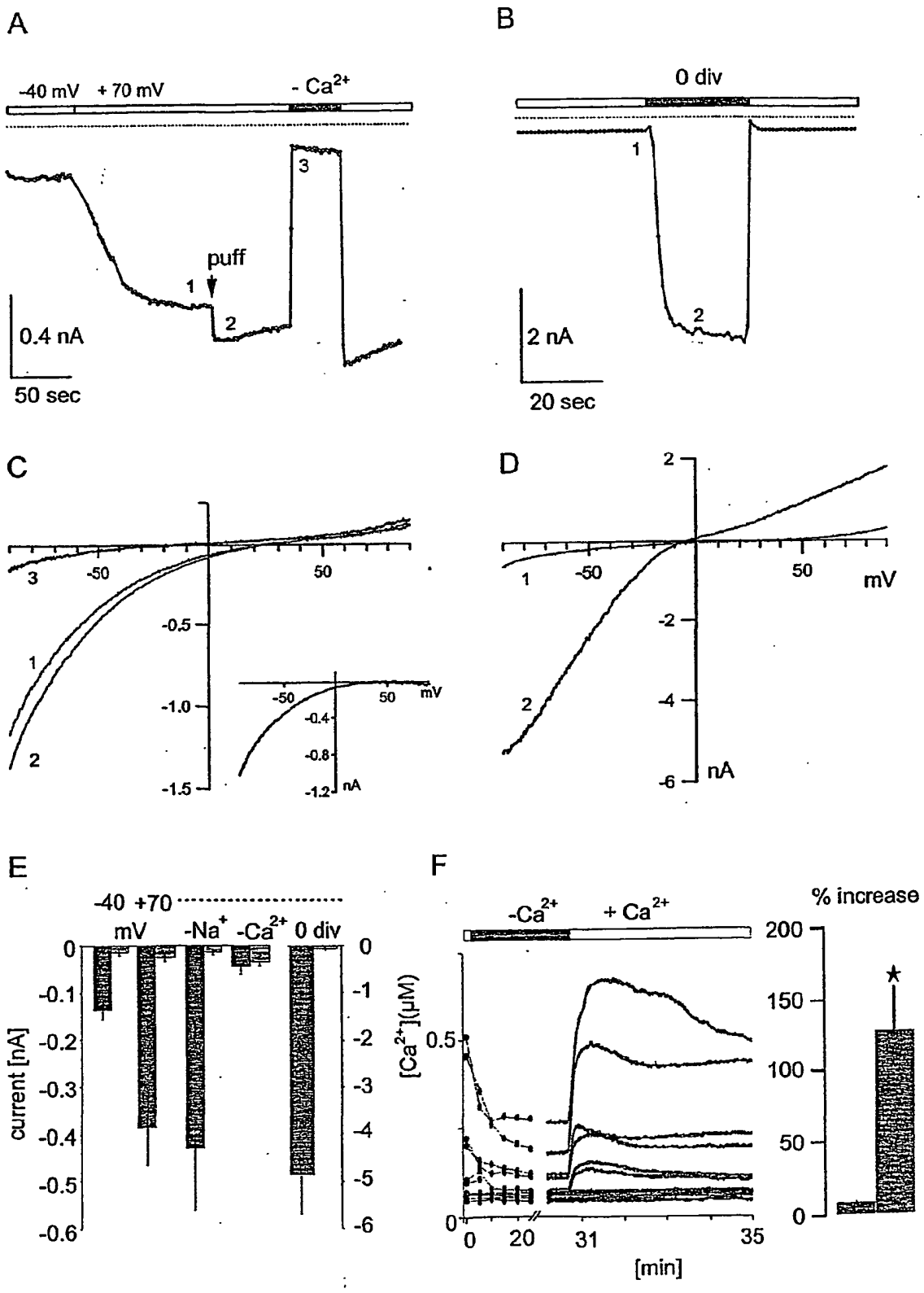

A
-40 mV
+ 70 mV
- Ca2+
puff
1
2
3
0.4 nA
50 sec
B
0 div
1
2
2 nA
20 sec
C
3
-50
50
1
2
-0.5
-1.0
-1.5
-50
50
mV
-0.4
-0.8
-1.2
nA
D
2
1
-50
50
mV
-2
-4
-6
nA
E
-40 +70
mV
-Na+
-Ca2+
0 div
current [nA]
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
0
-1
-2
-3
-4
-5
-6
F
-Ca2+
+ Ca2+
% increase
[Ca2+](µM)
0.5
0
0
20
31
35
[min]
200
150
100
50
0

Fig. 4

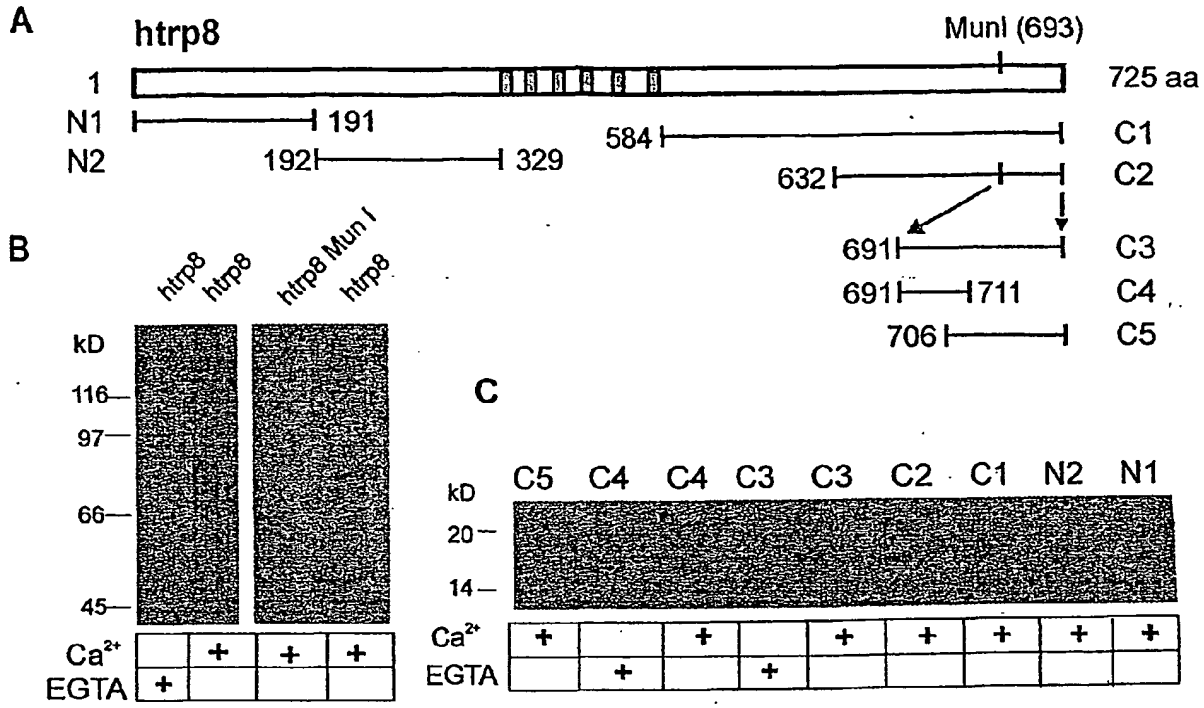

A
htrp8
MunI (693)
1
725 aa
N1
191
N2
192
329
584
C1
632
C2
691
C3
691
711
C4
706
C5
B
htrp8
htrp8
htrp8 Mun I
htrp8
kD
116
97
66
45
Ca2+
EGTA
C
C5 C4 C4 C3 C3 C2 C1 N2 N1
kD
20
14
Ca2+
EGTA

Fig. 5

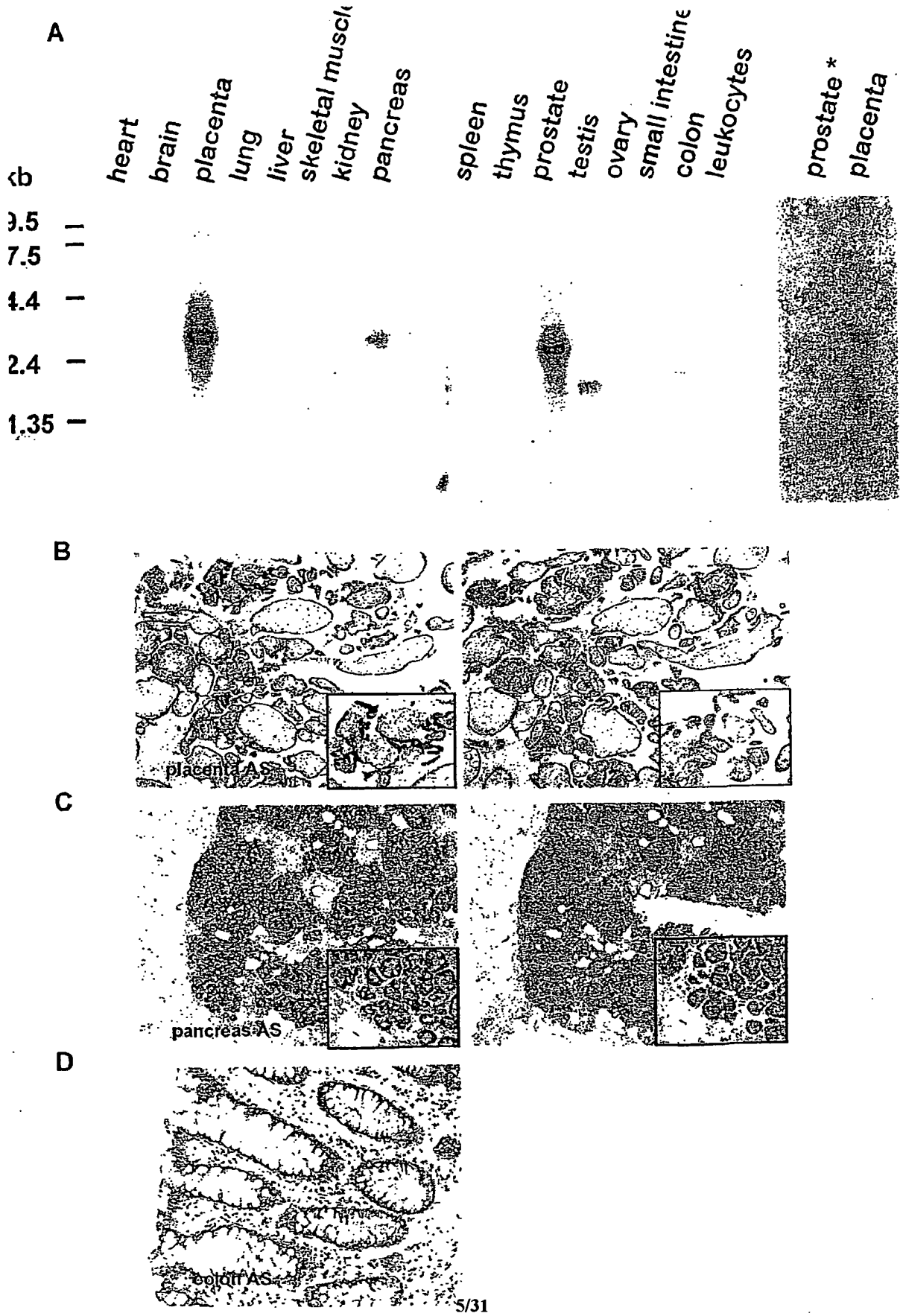

A
heart
brain
placenta
lung
liver
skeletal muscl
kidney
pancreas
spleen
thymus
prostate
testis
ovary
small intestine
colon
leukocytes
prostate *
placenta
kb
9.5
7.5
4.4
2.4
1.35
B
placenta AS
C
pancreas AS
D
colon AS

Fig. 6

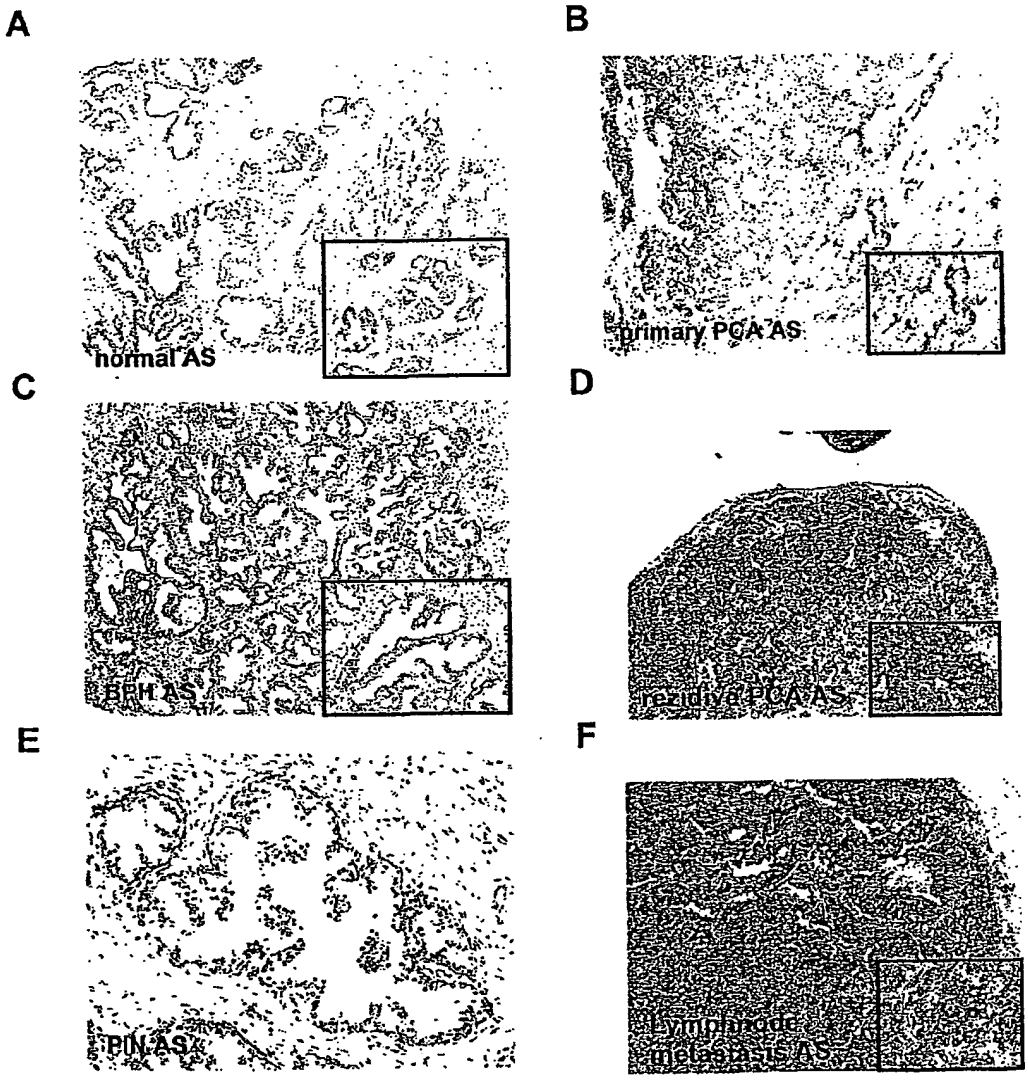

A
normal AS
B
primary PCA AS
C
BPH AS
D
rezidiva PCA AS
E
PIN AS
F
Lymphnode
metastasis AS

Fig. 7

```
       10                  30                  50
GCCAAGTGTAACAAACTCACAGCCCTCTCCAAACTGGCTGGGGCTGCTGGGAGACTCCCA
       70                  90                  110
AGGAACTCGTCAGGAAGGCAGGAGACAGGAGACGGGACCTCTACAGGGAGACGGTGGGCC
      130                 150                 170
GGCCCTTGGGGGGGCTGATGTGGCCCCAAGGCTGAGTCCCGTCAGGGTCTGGCCTCGGCC
      190                 210                 230
TCAGGCCCCCAAGGAGCCGGCCCTACACCCCATGGGTTTGTCACTGCCCAAGGAGAAAGG
                                  M  G  L  S  L  P  K  E  K  G
      250                 270                 290
GCTAATTCTCTGCCTATGGAGCAAGTTCTGCAGATGGTTCCAGAGACGGGAGTCCTGGGC
 L  I  L  C  L  W  S  K  F  C  R  W  F  Q  R  R  E  S  W  A
      310                 330                 350
CCAGAGCCGAGATGAGCAGAACCTGCTGCAGCAGAAGAGGATCTGGGAGTCTCCTCTCCT
 Q  S  R  D  E  Q  N  L  L  Q  Q  K  R  I  W  E  S  P  L  L
      370                 390                 410
TCTAGCTGCCAAAGATAATGATGTCCAGGCCCTGAACAAGTTGCTCAAGTATGAGGATTG
 L  A  A  K  D  N  D  V  Q  A  L  N  K  L  L  K  Y  E  D  C
      430                 450                 470
CAAGGTGCACCAGAGAGGAGCCATGGGGGAAACAGCGCTACACATAGCAGCCCTCTATGA
 K  V  H  Q  R  G  A  M  G  E  T  A  L  H  I  A  A  L  Y  D
      490                 510                 530
CAACCTGGAGGCCGCCATGGTGCTGATGGAGGCTGCCCCGGAGCTGGTCTTTGAGCCCAT
 N  L  E  A  A  M  V  L  M  E  A  A  P  E  L  V  F  E  P  M
      550                 570                 590
GACATCTGAGCTCTATGAGGGTCAGACTGCACTGCACATCGCTGTTGTGAACCAGAACAT
 T  S  E  L  Y  E  G  Q  T  A  L  H  I  A  V  V  N  Q  N  M
      610                 630                 650
GAACCTGGTGCGAGCCCTGCTTGCCCGCAGGGCCAGTGTCTCTGCCAGAGCCACAGGCAC
 N  L  V  R  A  L  L  A  R  R  A  S  V  S  A  R  A  T  G  T
      670                 690                 710
TGCCTTCCGCCGTAGTCCCCGCAACCTCATCTACTTTGGGGAGCACCCTTTGTCCTTTGC
 A  F  R  R  S  P  R  N  L  I  Y  F  G  E  H  P  L  S  F  A
      730                 750                 770
TGCCTGTGTGAACAGTGAGGAGATCGTGCGGCTGCTCATTGAGCATGGAGCTGACATCCG
 A  C  V  N  S  E  E  I  V  R  L  L  I  E  H  G  A  D  I  R
      790                 810                 830
GGCCCAGGACTCCCTGGGAAACACAGTGTTACACATCCTCATCCTCCAGCCCAACAAAAC
 A  Q  D  S  L  G  N  T  V  L  H  I  L  I  L  Q  P  N  K  T
      850                 870                 890
CTTTGCCTGCCAGATGTACAACCTGTTGCTGTCCTACGACAGACATGGGGACCACCTGCA
 F  A  C  Q  M  Y  N  L  L  L  S  Y  D  R  H  G  D  H  L  Q
      910                 930                 950
GCCCCTGGACCTCGTGCCCAATCACCAGGGTCTCACCCCTTTCAAGCTGGCTGGAGTGGA
 P  L  D  L  V  P  N  H  Q  G  L  T  P  F  K  L  A  G  V  E
      970                 990                 1010
GGGTAACACTGTGATGTTTCAGCACCTGATGCAGAAGCGGAAGCACACCCAGTGGACGTA
 G  N  T  V  M  F  Q  H  L  M  Q  K  R  K  H  T  Q  W  T  Y
     1030                1050                1070
TGGACCACTGACCTCGACTCTCTATGACCTCACAGAGATCGACTCCTCAGGGGATGAGCA
 G  P  L  T  S  T  L  Y  D  L  T  E  I  D  S  S  G  D  E  Q
     1090                1110                1130
GTCCCTGCTGGAACTTATCATCACCACCAAGAAGCGGGAGGCTCGCCAGATCCTGGACCA
 S  L  L  E  L  I  I  T  T  K  K  R  E  A  R  Q  I  L  D  Q
     1150                1170                1190
GACGCCGGTGAAGGAGCTGGTGAGCCTCAAGTGGAAGCGGTACGGGCGGCCGTACTTCTG
 T  P  V  K  E  L  V  S  L  K  W  K  R  Y  G  R  P  Y  F  C
     1210                1230                1250
CATGCTGGGTGCCATATATCTGCTGTACATCATCTGCTTCACCATGTGCTGCATCTACCG
 M  L  G  A  I  Y  L  L  Y  I  I  C  F  T  M  C  C  I  Y  R
     1270                1290                1310
```

Fig. 7 / continuation 1

```
CCCCCTCAAGCCCAGGACCAATAACCGCACAAGCCCCCGGGACAACACCCTCTTACAGCA
 P  L  K  P  R  T  N  N  R  T  S  P  R  D  N  T  L  L  Q  Q
      1330                1350                1370
GAAGCTACTTCAGGAAGCCTACGTGACCCCTAAGGACGATATCCGGCTGGTCGGGGAGCT
 K  L  L  Q  E  A  Y  V  T  P  K  D  D  I  R  L  V  G  E  L
      1390                1410                1430
GGTGACTGTCATTGGGGCTATCATCATCCTGCTGGTAGAGGTTCCAGACATCTTCAGAAT
 V  T  V  I  G  A  I  I  I  L  L  V  E  V  P  D  I  F  R  M
      1450                1470                1490
GGGGGTCACTCGCTTCTTTGGACAGACCATCCTTGGGGGCCCATTCCATGTCCTCATCAT
 G  V  T  R  F  F  G  Q  T  I  L  G  G  P  F  H  V  L  I  I
      1510                1530                1550
CACCTATGCCTTCATGGTGCTGGTGACCATGGTGATGCGGCTCATCAGTGCCAGCGGGGA
 T  Y  A  F  M  V  L  V  T  M  V  M  R  L  I  S  A  S  G  E
      1570                1590                1610
GGTGGTACCCATGTCCTTTGCACTCGTGCTGGGCTGGTGCAACGTCATGTACTTCGCCCG
 V  V  P  M  S  F  A  L  V  L  G  W  C  N  V  M  Y  F  A  R
      1630                1650                1670
AGGATTCCAGATGCTAGGCCCCTTCACCATCATGATTCAGAAGATGATTTTTGGCGACCT
 G  F  Q  M  L  G  P  F  T  I  M  I  Q  K  M  I  F  G  D  L
      1690                1710                1730
GATGCGATTCTGCTGGCTGATGGCTGTGGTCATCCTGGGCTTTGCTTCAGCCTTCTATAT
 M  R  F  C  W  L  M  A  V  V  I  L  G  F  A  S  A  F  Y  I
      1750                1770                1790
CATCTTCCAGACAGAGGACCCCGAGGAGCTAGGCCACTTCTACGACTACCCCATGGCCCT
 I  F  Q  T  E  D  P  E  E  L  G  H  F  Y  D  Y  P  M  A  L
      1810                1830                1850
GTTCAGCACCTTCGAGCTGTTCCTTACCATCATCGATGGCCCAGCCAACTACAACGTGGA
 F  S  T  F  E  L  F  L  T  I  I  D  G  P  A  N  Y  N  V  D
      1870                1890                1910
CCTGCCCTTCATGTACAGCATCACCTATGCTGCCTTTGCCATCATCGCCACACTGCTCAT
 L  P  F  M  Y  S  I  T  Y  A  A  F  A  I  I  A  T  L  L  M
      1930                1950                1970
GCTCAACCTCCTCATTGCCATGATGGGCGACACTCACTGGCGAGTGGCCCATGAGCGGGA
 L  N  L  L  I  A  M  M  G  D  T  H  W  R  V  A  H  E  R  D
      1990                2010                2030
TGAGCTGTGGAGGGCCCAGATTGTGGCCACCACGGTGATGCTGGAGCGGAAGCTGCCTCG
 E  L  W  R  A  Q  I  V  A  T  T  V  M  L  E  R  K  L  P  R
      2050                2070                2090
CTGCCTGTGGCCTCGCTCCGGGATCTGCGGACGGGAGTATGGCCTGGGGGACCGCTGGTT
 C  L  W  P  R  S  G  I  C  G  R  E  Y  G  L  G  D  R  W  F
      2110                2130                2150
CCTGCGGGTGGAAGACAGGCAAGATCTCAACCGGCAGCGGATCCAACGCTACGCACAGGC
 L  R  V  E  D  R  Q  D  L  N  R  Q  R  I  Q  R  Y  A  Q  A
      2170                2190                2210
CTTCCACACCCGGGGCTCTGAGGATTTGGACAAAGACTCAGTGGAAAAACTAGAGCTGGG
 F  H  T  R  G  S  E  D  L  D  K  D  S  V  E  K  L  E  L  G
      2230                2250                2270
CTGTCCCTTCAGCCCCCACCTGTCCCTTCCTACGCCCTCAGTGTCTCGAAGTACCTCCCG
 C  P  F  S  P  H  L  S  L  P  T  P  S  V  S  R  S  T  S  R
      2290                2310                2330
CAGCAGTGCCAATTGGGAAAGGCTTCGGCAAGGGACCCTGAGGAGAGACCTGCGTGGGAT
 S  S  A  N  W  E  R  L  R  Q  G  T  L  R  R  D  L  R  G  I
      2350                2370                2390
AATCAACAGGGGTCTGGAGGACGGGGAGAGCTGGGAATATCAGATCTGACTGCGTGTTCT
 I  N  R  G  L  E  D  G  E  S  W  E  Y  Q  I
      2410                2430                2450
CACTTCGCTTCCTGGAACTTGCTCTCATTTTCCTGGGTGCATCAAACAAAACAAAAACCA
      2470                2490                2510
AACACCCAGAGGTCTCATCTCCCAGGCCCCAGGGAGAAAGAGGAGTAGCATGAACGCCAA
      2530                2550                2570
GGAATGTACGTTGAGAATCACTGCTCCAGGCCTGCATTACTCCTTCAGCTCTGGGGCAGA
```

Fig. 7 / continuation 2

```
      2590                2610                2630
GGAAGCCCAGCCCAAGCACGGGGCTGGCAGGGCGTGAGGAACTCTCCTGTGGCCTGCTCA
      2650                2670                2690
TCACCCTTCCGACAGGAGCACTGCATGTCAGAGCACTTTAAAAACAGGCCAGCCTGCTTG
      2710                2730                2750
GGCCCTCGGTCTCCACCCCAGGGTCATAAGTGGGGAGAGAGCCCTTCCCACGGCACCCAG
      2770                2790                2810
GCAGGTGCAGGGAAGTGCAGAGCTTGTGGAAAGCGTGTGAGTGAGGGAGACAGGAACGGC
      2830                2850                2870
TCTGGGGGTGGGAAGTGGGGCTAGGTCTTGCCAACTCCATCTTCAATAAAGTCGTTTTCG
      2890                2910
GATCCCTAAAAAAAAAAAAAAAAAAAAAAAAAA
```

```
MGLSLPKEKGLILCLWSKFCRWFQRRESWAQSRDEQNLLQQKRIWESPLLLAAKDNDVQALNKLLKYEDCKVHQRGAMGETALHIA
ALYDNLEAAMVLMEAAPELVFEPMTSELYEGQTALHIAVVNQNMNLVRALLARRASVSARATGTAFRRSPRNLIYFGEHPLSFAAC
VNSEEIVRLLIEHGADIRAQDSLGNTVLHILILQPNKTFACQMYNLLLSYDRHGDHLQPLDLVPNHQGLTPFKLAGVEGNTVMFQH
LMQKRKHTQWTYGPLTSTLYDLTEIDSSGDEQSLLELIITTKKREARQILDQTPVKELVSLKWKRYGRPYFCMLGAIYLLYIICFT
MCCIYRPLKPRTNNRTSPRDNTLLQQKLLQEAYVTPKDDIRLVGELVTVIGAIIILLVEVPDIFRMGVTRFFGQTILGGPFHVLII
TYAFMVLVTMVMRLISASGEVVPMSFALVLGWCNVMYFARGFQMLGPFTIMIQKMIFGDLMRFCWLMAVVILGFASAFYIIFQTED
PEELGHFYDYPMALFSTFELFLTIIDGPANYNVDLPFMYSITYAAFAIIATLLMLNLLIAMMGDTHWRVAHERDELWRAQIVATTV
MLERKLPRCLWPRSGICGREYGLGDRWFLRVEDRQDLNRQRIQRYAQAFHTRGSEDLDKDSVEKLELGCPFSPHLSLPTPSVSRST
SRSSANWERLRQGTLRRDLRGIINRGLEDGESWEYQI
```

Figure 8:

A)

```
                                ATGGGTTTGTCACTGCCCAAGGAGAAAGGGCTAATTCTCT
                                M  G  L  S  L  P  K  E  K  G  L  I  L  C
        250                     270                     290
GCCTATGGAGCAAGTTCTGCAGATGGTTCCAGAGACGGGAGTCCTGGGCCCAGAGCCGAG
  L  W  S  K  F  C  R  W  F  Q  R  R  E  S  W  A  Q  S  R  D
        310                     330                     350
ATGAGCAGAACCTGCTGCAGCAGAAGAGGATCTGGGAGTCTCCTCTCCTTCTAGCTGCCA
  E  Q  N  L  L  Q  Q  K  R  I  W  E  S  P  L  L  L  A  A  K
        370                     390                     410
AAGATAATGATGTCCAGGCCCTGAACAAGTTGCTCAAGTATGAGGATTGCAAGGTGCACC
  D  N  D  V  Q  A  L  N  K  L  L  K  Y  E  D  C  K  V  H  Q
        430                     450                     470
AGAGAGGAGCCATGGGGGAAACAGCGCTACACATAGCAGCCCTCTATGACAACCTGGAGG
  R  G  A  M  G  E  T  A  L  H  I  A  A  L  Y  D  N  L  E  A
        490                     510                     530
CCGCCATGGTGCTGATGGAGGCTGCCCCGGAGCTGGTCTTTGAGCCCATGACATCTGAGC
  A  M  V  L  M  E  A  A  P  E  L  V  F  E  P  M  T  S  E  L
        550                     570                     590
TCTATGAGGGTCAGACTGCACTGCACATCGCTGTTGTGAACCAGAACATGAACCTGGTGC
  Y  E  G  Q  T  A  L  H  I  A  V  V  N  Q  N  M  N  L  V  R
        610                     630                     650
GAGCCCTGCTTGCCCGCAGGGCCAGTGTCTCTGCCAGAGCCACAGGCACTGCCTTCCGCC
  A  L  L  A  R  R  A  S  V  S  A  R  A  T  G  T  A  F  R  R
        670                     690                     710
GTAGTCCCTGCAACCTCATCTACTTTGGGGAGCACCCTTTGTCCTTTGCTGCCTGTGTGA
  S  P  C  N  L  I  Y  F  G  E  H  P  L  S  F  A  A  C  V  N
```

Fig. 8 / contin ı 1

```
      730                 750                 770
ACAGTGAGGAGATCGTGCGGCTGCTCATTGAGCATGGAGCTGACATCCGGGCCCAGGACT
  S  E  E  I  V  R  L  L  I  E  H  G  A  D  I  R  A  Q  D  S
      790                 810                 830
CCCTGGGAAACACAGTGTTACACATCCTCATCCTCCAGCCCAACAAAACCTTTGCCTGCC
  L  G  N  T  V  L  H  I  L  I  L  Q  P  N  K  T  F  A  C  Q
      850                 870                 890
AGATGTACAACCTGTTGCTGTCCTACGACAGACATGGGGACCACCTGCAGCCCCTGGACC
  M  Y  N  L  L  L  S  Y  D  R  H  G  D  H  L  Q  P  L  D  L
      910                 930                 950
TCGTGCCCAATCACCAGGGTCTCACCCCTTTCAAGCTGGCTGGAGTGGAGGGTAACACTG
  V  P  N  H  Q  G  L  T  P  F  K  L  A  G  V  E  G  N  T  V
      970                 990                1010
TGATGTTTCAGCACCTGATGCAGAAGCGGAAGCACACCCAGTGGACGTATGGACCACTGA
  M  F  Q  H  L  M  Q  K  R  K  H  T  Q  W  T  Y  G  P  L  T
     1030                1050                1070
CCTCGACTCTCTATGACCTCACAGAGATCGACTCCTCAGGGGATGAGCAGTCCCTGCTGG
  S  T  L  Y  D  L  T  E  I  D  S  S  G  D  E  Q  S  L  L  E
     1090                1110                1130
AACTTATCATCACCACCAAGAAGCGGGAGGCTCGCCAGATCCTGGACCAGACGCCGGTGA
  L  I  I  T  T  K  K  R  E  A  R  Q  I  L  D  Q  T  P  V  K
     1150                1170                1190
AGGAGCTGGTGAGCCTCAAGTGGAAGCGGTACGGGCGGCCGTACTTCTGCATGCTGGGTG
  E  L  V  S  L  K  W  K  R  Y  G  R  P  Y  F  C  M  L  G  A
     1210                1230                1250
CCATATATCTGCTGTACATCATCTGCTTCACCATGTGCTGCATCTACCGCCCCCTCAAGC
  I  Y  L  L  Y  I  I  C  F  T  N  C  C  I  Y  R  P  L  K  P
     1270                1290                1310
CCAGGACCAATAACCGCACGAGCCCCCGGGACAACACCCTCTTACAGCAGAAGCTACTTC
  R  T  N  N  R  T  S  P  R  D  N  T  L  L  Q  Q  K  L  L  Q
     1330                1350                1370
AGGAAGCCTACATGACCCCTAAGGACGATATCCGGCTGGTCGGGGAGCTGGTGACTGTCA
  E  A  Y  M  T  P  K  D  D  I  R  L  V  G  E  L  V  T  V  I
     1390                1410                1430
TTGGGGCTATCATCATCCTGCTGGTAGAGGTTCCAGACATCTTCAGAATGGGGGTCACTC
  G  A  I  I  I  L  L  V  E  V  P  D  I  F  R  M  G  V  T  R
     1450                1470                1490
GCTTCTTTGGACAGACCATCCTTGGGGGCCCATTCCATGTCCTCATCATCACCTATGCCT
  F  F  G  Q  T  I  L  G  G  P  F  H  V  L  I  I  T  Y  A  F
     1510                1530                1550
TCATGGTGCTGGTGACCATGGTGATGCGGCTCATCAGTGCCAGCGGGGAGGTGGTACCCA
  M  V  L  V  T  M  V  M  R  L  I  S  A  S  G  E  V  V  P  M
     1570                1590                1610
TGTCCTTTGCACTCGTGCTGGGCTGGTGCAACGTCATGTACTTCGCCCGAGGATTCCAGA
  S  F  A  L  V  L  G  W  C  N  V  M  Y  F  A  R  G  F  Q  M
     1630                1650                1670
TGCTAGGCCCCTTCACCATCATGATTCAGAAGATGATTTTTGGCGACCTGATGCGATTCT
  L  G  P  F  T  I  M  I  Q  K  M  I  F  G  D  L  M  R  F  C
     1690                1710                1730
GCTGGCTGATGGCTGTGGTCATCCTGGGCTTTGCTTCAGCCTTCTATATCATCTTCCAGA
  W  L  M  A  V  V  I  L  G  F  A  S  A  F  Y  I  I  F  Q  T
     1750                1770                1790
CAGAGGACCCCGAGGAGCTAGGCCACTTCTACGACTACCCCATGGCCCTGTTCAGCACCT
  E  D  P  E  E  L  G  H  F  Y  D  Y  P  M  A  L  F  S  T  F
     1810                1830                1850
TCGAGCTGTTCCTTACCATCATCGATGGCCCAGCCAACTACAACGTGGACCTGCCCTTCA
  E  L  F  L  T  I  I  D  G  P  A  N  Y  N  V  D  L  P  F  M
     1870                1890                1910
TGTACAGCATCACCTATGCTGCCTTTGCCATCATCGCCACACTGCTCATGCTCAACCTCC
  Y  S  I  T  Y  A  A  F  A  I  I  A  T  L  L  M  L  N  L  L
     1930                1950                1970
TCATTGCCATGATGGGCGACACTCACTGGCGAGTGGCCCATGAGCGGGATGAGCTGTGGA
```

Fig. 8 / continuation 2

```
  I  A  M  M  G  D  T  H  W  R  V  A  H  E  R  D  E  L  W  R
     1990              2010              2030
GGGCCCAGATTGTGGCCACCACGGTGATGCTGGAGCGGAAGCTGCCTCGCTGCCTGTGGC
  A  Q  I  V  A  T  T  V  M  L  E  R  K  L  P  R  C  L  W  P
     2050              2070              2090
CTCGCTCCGGGATCTGCGGACGGGAGTATGGCCTGGGAGACCGCTGGTTCCTGCGGGTGG
  R  S  G  I  C  G  R  E  Y  G  L  G  D  R  W  F  L  R  V  E
     2110              2130              2150
AAGACAGGCAAGATCTCAACCGGCAGCGGATCCAACGCTACGCACAGGCCTTCCACACCC
  D  R  Q  D  L  N  R  Q  R  I  Q  R  Y  A  Q  A  F  H  T  R
     2170              2190              2210
GGGGCTCTGAGGATTTGGACAAAGACTCAGTGGAAAAACTAGAGCTGGGCTGTCCCTTCA
  G  S  E  D  L  D  K  D  S  V  E  K  L  E  L  G  C  P  F  S
     2230              2250              2270
GCCCCCACCTGTCCCTTCCTATGCCCTCAGTGTCTCGAAGTACCTCCCGCAGCAGTGCCA
  P  H  L  S  L  P  M  P  S  V  S  R  S  T  S  R  S  S  A  N
     2290              2310              2330
ATTGGGAAAGGCTTCGGCAAGGGACCCTGAGGAGAGACCTGCGTGGGATAATCAACAGGG
  W  E  R  L  R  Q  G  T  L  R  R  D  L  R  G  I  I  N  R  G
     2350              2370              2390
GTCTGGAGGACGGGGAGAGCTGGGAATATCAGATCTGA
  L  E  D  G  E  S  W  E  Y  Q  I  *
```

```
MGLSLPKEKGLILCLWSKFCRWFQRRESWAQSRDEQNLLQQKRIWESPLLLAAKDNDVQALNKLLKYEDCKVHQRGAMGETALHIA
ALYDNLEAAMVLMEAAPELVFEPMTSELYEGQTALHIAVVNQNMNLVRALLARRASVSARATGTAFRRSPCNLIYFGEHPLSFAAC
VNSEEIVRLLIEHGADIRAQDSLGNTVLHILILQPNKTFACQMYNLLLSYDRHGDHLQPLDLVPNHQGLTPFKLAGVEGNTVMFQH
LMQKRKHTQWTYGPLTSTLYDLTEIDSSGDEQSLLELIITTKKREARQILDQTPVKELVSLKWKRYGRPYFCMLGAIYLLYIICFT
MCCIYRPLKPRTNNRTSPRDNTLLQQKLLQEAYMTPKDDIRLVGELVTVIGAIIILLVEVPDIFRMGVTRFFGQTILGGPFHVLII
TYAFMVLVTNVMRLISASGEVVPMSFALVLGWCNVMYFARGFQMLGPFTIMIQKMIFGDLMRFCWLMAVVILGFASAFYIIFQTED
PEELGHFYDYPMALFSTFELFLTIIDGPANYNVDLPFMYSITYAAFAIIATLLMLNLLIAMMGDTHWRVAHERDELWRAQIVATTV
MLERKLPRCLWPRSGICGREYGLGDRWFLRVEDRQDLNRQRIQRYAQAFHTRGSEDLDKDSVEKLELGCPFSPHLSLPMPSVSRST
SRSSANWERLRQGTLRRDLRGIINRGLEDGESWEYQI
```

B)

```
CAAACTCACAGCCCTCTCCAAACTGGCTGGGGCTGCTGGGAGACTCCCAAGGAACTCGTCAGGAAGGCAGGAGACAGGAGACGGGA
CCTCTACAGGGAGACGGTGGGCCGGCCCTTGGGGGGGCTGATGTGGCCCCAAGGCTGAGTCCCGTCAGGGTCTGGCCTCGGCCTCA
GGCCCCCAAGGAGCCGGCCCTACACCCCATGGGTTTGTCACTGCCCAAGGAGAAAGGCTAATTCTCTGCCTATGGAGCAAGTTCT
GCAGATGGTTCCAGAGACGGGAGTCCTGGGCCCAGAGCCGAGATGAGCAGAACCTGCTGCAGCAGAAGAGGATCTGGGAGTCTCCT
CTCCTTCTAGCTGCCAAAGATAATGATGTCCAGGCCCTGAACAAGTTGCTCAAGTATGAGGATTGCAAGGTGCACCAGAGAGGAGC
CATGGGGGAAACAGCGCTACACATAGCAGCCCTCTATGACAACCTGGAGGCCGCCATGGTGCTGATGGAGGCTGCCCCGGAGCTGG
TCTTTGAGCCCATGACATCTGAGCTCTATGAGGGTCAGACTGCACTGCACATCGCTGTTGTGAACCAGAACATGAACCTGGTGCGA
GCCCTGCTTGCCCGCAGGGCCAGTGTCTCTGCCAGAGCCACAGGCACTGCCTTCCGCCGTAGTCCCCGCAACCTCATCTACTTTGG
GGAGCACCCTTTGTCCTTTGCTGCCTGTGTGAACAGTGAGGAGATCGTCGCGCTGCTCATTGAGCATGGAGCTGACATCCGGGCCC
AGGACTGCCTGGCCCAACAAAACCTTTGCCTGCCAGATGTACAACCTGTTGCTGTCCTACGACAGACATGGGGACCACCTGCAGCC
CCTGGACCTCGTGCCCAATCACCAGGGTCTCACCCCTTTCAAGCTGGCTGGAGTGGAGGGTAACACTGTGATGTTTCAGCACCTGA
TGCAGAAGCGGAAGCACACCCAGTGGACGTATGGACCACTGACCTCGACTCTCTATGACCTCACAGAGATCGACTCCTCAGGGGAT
GAGCAGTCCCTGCTGGAACTTATCATCACCACCAAGAAGCGGGAGGCTCGCCAGATCCTGGACCAGACGCCGGTGAAGGAGCTGGT
GAGCCTCAAGTGGAAGCGGTACGGGCGGCCGTACTTCTGCATGCTGGGTGCCATATATCTGCTGTACATCATCTGCTTCACCATGT
GCTGCATCTACCGCCCCCTCAAGCCCAGGACCAATAACCGCACAAGCCCCCGGGACAACACCCTCTTACAGCAGAAGCTACTTCAG
GAAGCCTACGTGACCCCTAAGGACGATATCCGGCTGGTCGGGGAGCTGGTGACTGTCATTGGGGCTATCATCATCCTGCTGGTAGA
GGTTCCAGACATCTTCAGAATGGGGGTCACTCGCTTCTTTGGACAGACCATCCTTGGGGGCCCATTCCATGTCCTCATCATCACCT
ATGCCTTCATGGTGCTGGTGACCATGGTGATGCGGCTCATATGATTTTGGCGACCTGATGCGATTCTGCTGGCTGATGGCTGTGG
TCATCCTGGGCTTTGCTTCAGCCTTCTATATCATCTTCCAGACAGAGGACCCCGAGGAGCTAGGCCACTTCTACGACTACCCCATG
GCCCTGTTCAGCACCTTCGAGCTGGTCCTTACCATCATCGATGGCCCAGCCAACTACAACGTGGACCTGCCCTTCATGTACAGCAT
CACCTATGCTGCCTTTGCCATCATCGCCACACTGCTCATGCTCAACCTCCTCATTGCCATGATGGGCGACACTCACTGGCGAGTGG
CCCATGAGCGGGATGAGCTGTGGAGGGCCCAGATTGTGGCCACCACGGTGATGCTGGAGCGGAAGCTGCCTCGCTGCCTGTGGCCT
CGCTCCGGGATCTGCGGACGGGAGTATGGCCTGGGGGACCGCTGGTTCCTGCGGGTGGAAGACAGGCAAGATCTCAACCGGCAGCG
```

Fig. 8 / continuation 3

GATCCAACGCTACGCACAGGCCTTCCACACCCGGGGCTCTGAGGATTTGGACAAAGACTCAGTGGAAAAACTAGAGCTGGGCTGTC
CCTTCAGCCCCCACCTGTCCCTTCCTACGCCCTCAGTGTCTCGAAGTACCTCCCGCAGCAGTGCCAATTGGGAAAGGCTTCGGCAA
GGGACCCTGAGGAGAGACCTGCGTGGGATAATCAACAGGGGTCTGGAGGACGGGGAGAGCTGGGAATATCAGATCTGACTGCGTGT
TCTCACTTCGCTTCCTGGAACTTGCTCTCATTTTCCTGGGTGCATCAAACAAAACAAAAACCAAACACCCAGAGGTCTCATCTCCC
AGGCCCCAGGGAGAAAGAGGAGTAGCATGAACGCCAAGGAATGTACGTTGAGAATCACTGCTCCAGGCCTGCATTACTCCTTCAGC
TCTGGGGCAGAGGAAGCCCAGCCCAAGCACGGGGCTGGCAGGGCGTGAGGAACTCTCCTGTGGCCTGCTCATCACCCTTCCGACAG
GAGCACTGCATGTCAGAGCACTTTAAAAACAGGCCAGCCTGCTTGGGCCCTCGGTCTCCACCCCAGGGTCATAAGTGGGGAGAGAG
CCCTTCCCAGGGCACCCAGGCAGGTGCAGGGAAGTGCAGAGCTTGTGGAAAGCGTGTGAGTGAGGGAGACAGGAACGGCTCTGGGG
GTGGGAAGTGGGGCTAGGTCTTGCCAACTCCATCTTCAATAAAGTCGTTTTCGGATCCCTAAAAAAAAAAAAAAAAAAAAAAAAAA

c.)

CAAACTCACAGCCCTCTCCAAACTGGCTGGGGCTGCTGGGAGACTCCCAAGGAACTCGTCAGGAAGGCAGGAGACAGGAGACGGGA
CCTCTACAGGGAGACGGTGGGCCGGCCCTTGGGGGGGCTGATGTGGCCCCAAGGCTGAGTCCCGTCAGGGTCTGGCCTCGGCCTCA
GGCCCCCAAGGAGCCGGCCCTACACCCCATGGGTTTGTCACTGCCCAAGGAGAAAGGGCTAATTCTCTGCCTATGGAGCAAGTTCT
GCAGATGGTTCCAGAGACGGGAGTCCTGGGCCCAGAGCCGAGATGAGCAGAACCTGCTGCAGCAGAAGAGGATCTGGGAGTCTCCT
CTCCTTCTAGCTGCCAAAGATAATGATGTCCAGGCCCTGAACAAGTTGCTCAAGTATGAGGATTGCAAGGTGCACCAGAGAGGAGC
CATGGGGGAAACAGCGCTACACATAGCAGCCCTCTATGACAACCTGGAGGCCGCCATGGTGCTGATGGAGGCTGCCCCGGAGCTGG
TCTTTGAGCCCATGACATCTGAGCTCTATGAGGTCCTGACTGCCCATCACTTGAACGCCTGCCCCCTGAAATGCCAGGGCCTAGAG
AAGAGGAAGAGATGGGCAGCAGCTGGATCCCCTGGGAATCCTGAACACCCGAGAGCTCCCTGTTCTCCATCCCAGGCTACCCCTGA
GGGAAAGAGACTGGGGTGCATATGGGAGGGACCCCCTGCAGGATCCTGGGGACAGACCCGTGACTGACAGCTGTCTCTGGGCCAGG
TCAGACTGCACTGCACATCGCTGTTGTGAACCAGAACATGAACCTGGTGCGAGCCCTGCTTGCCCGCAGGGCCAGTGTCTCTGCCA
GAGCCACAGGCACTGCCTTCCGCCGTAGTCCCTGCAACCTCATCTACTTTGGGGAGCACCCTTTGTCCTTTGCTGCCTGTGTGAAC
AGTGAGGAGATCGTGCGGCTGCTCATTGAGCATGGAGCTGACATCCGGGCCCAGGACTCCCTGGCCCAACAAAACCTTTGCCTGCC
AGATGTACAACCTGTTGCTGTCCTACGACAGACATGGGGACCACCTGCAGCCCCTGGACCTCGTGCCCAATCACCAGGGTCTCACC
CCTTTCAAGCTGGCTGGAGTGGAGGGTAACACTGTGATGTTTCAGCACCTGATGCAGAAGCGGAAGCACACCCAGTGGACGTATGG
ACCACTGACCTCGACTCTCTATGACCTCACAGAGATCGACTCCTCAGGGGATGAGCAGTCCCTGCTGGAACTTATCATCACCACCA
AGAAGCGGGAGGCTCGCCAGATCCTGGACCAGACGCCGGTGAAGGAGCTGGTGAGCCTCAAGTGGAAGCGGTACGGGCGGCCGTAC
TTCTGCATGCTGGGTGCCATATATCTGCTGTACATCATCTGCTTCACCATGTGCTGCATCTACCGCCCCCTCAAGCCCAGGACCAA
TAACCGCACGAGCCCCCGGGACAACACCCTCTTACAGCAGAAGCTACTTCAGGAAGCCTACATGACCCCTAAGGACGATATCCGGC
TGGTCGGGGAGCTGGTGACTGTCATTGGGGCTATCATCATCCTGCTGGTAGAGGTTCCAGACATCTTCAGAATGGGGGTCACTCGC
TTCTTTGGACAGACCATCCTTGGGGGCCCATTCCATGTCCTCATCATCACCTATGCCTTCATGGTGCTGGTGACCATGGTGATGCG
GCTCATCAGTGCCAGCGGGGAGGTGGTACCCATGTCCTTTGCACTCGTGCTGGGCTGGTGCAACGTCATGTACTTCGCCCGAGGAT
TCCAGATGCTAGGCCCCTTCACCATCATGATTCAGAAGATGATTTTTGGCGACCTGATGCGATTCTGCTGGCTGATGGCTGTGGTC
ATCCTGGGCTTTGCTTAGACAGAGGACCCCGAGGAGCTAGGCCACTTCTACGACTACCCCATGGCCCTGTTCAGCACCTTCGAGCT
GGTCCTTACCATCATCGATGGCCCAGCCAACTACAACGTGGACCTGCCCTTCATGTACAGCATCACCTATGCTGCCTTTGCCATCA
TCGCCACACTGCTCATGCTCAACCTCCTCATTGCCATGATGGGCGACACTCACTGGCGAGTGGCCCATGAGCGGGATGAGCTGTGG
AGGGCCCAGATTGTGGCCACCACGGTGATGCTGGAGCGGAAGCTGCCTCGCTGCCTGTGGCCTCGCTCCGGGATCTGCGGACGGGA
GTATGGCCTGGGAGACCGCTGGTTCCTGCGGGTGGAAGACAGGCAAGATCTCAACCGGCAGCGGATCCAACGCTACGCACAGGCCT
TCCACACCCGGGGCTCTGAGGATTTGGACAAAGACTCAGTGGAAAAACTAGAGCTGGGCTGTCCCTTCAGCCCCCACCTGTCCCTT
CCTATGCCCTCAGTGTCTCGAAGTACCTCCCGCAGCAGTGCCAATTGGGAAAGGCTTCGGCAAGGGACCCTGAGGAGAGACCTGCG
TGGGATAATCAACAGGGGTCTGGAGGACGGGGAGAGCTGGGAATATCAGATCTGACTGCGTGTTCTCACTTCGCTTCCTGGAACTT
GCTCTCATTTTCCTGGGTGCATCAAACAAAACAAAAACCAAACACCCAGAGGTCTCATCTCCCAGGCCCCAGGGAGAAAGAGGAGT
AGCATGAACGCCAAGGAATGTACGTTGAGAATCACTGCTCCAGGCCTGCATTACTCCTTCAGCTCTGGGGCAGAGGAAGCCCAGCC
CAAGCACGGGGCTGGCAGGGCGTGAGGAACTCTCCTGTGGCCTGCTCATCACCCTTCCGACAGGAGCACTGCATGTCAGAGCACTT
TAAAAACAGGCCAGCCTGCTTGGGCCCTCGGTCTCCACCCCAGGGTCATAAGTGGGGAGAGAGCCCTTCCCAGGGCACCCAGGCAG
GTGCAGGGAAGTGCAGAGCTTGTGGAAAGCGTGTGAGTGAGGGAGACAGGAACGGCTCTGGGGGTGGGAAGTGGGGCTAGGTCTTG
CCAACTCCATCTTCAATAAAGTCGTTTTCGGATCCCTAAAAAAAAAAAAAAAAAAAAAAAAAA

D.)

CAAACTCACAGCCCTCTCCAAACTGGCTGGGGCTGCTGGGAGACTCCCAAGGAACTCGTCAGGAAGGCAGGAGACAGGAGACGGGA
CCTCTACAGGGAGACGGTGGGCCGGCCCTTGGGGGGGCTGATGTGGCCCCAAGGCTGAGTCCCGTCAGGGTCTGGCCTCGGCCTCA
GGCCCCCAAGGAGCCGGCCCTACACCCCATGGGTTTGTCACTGCCCAAGGAGAAAGGGCTAATTCTCTGCCTATGGAGCAAGTTCT
GCAGATGGTTCCAGAGACGGGAGTCCTGGGCCCAGAGCCGAGATGAGCAGAACCTGCTGCAGCAGAAGAGGATCTGGGAGTCTCCT
CTCCTTCTAGCTGCCAAAGATAATGATGTCCAGGCCCTGAACAAGTTGCTCAAGTATGAGGATTGCAAGGTGCACCAGAGAGGAGC
CATGGGGGAAACAGCGCTACACATAGCAGCCCTCTATGACAACCTGGAGGCCGCCATGGTGCTGATGGAGGCTGCCCCGGAGCTGG
TCTTTGAGCCCATGACATCTGAGCTCTATGAGGGTCAGACTGCACTGCACATCGCTGTTGTGAACCAGAACATGAACCTGGTGCGA
GCCCTGCTTGCCCGCAGGGCCAGTGTCTCTGCCAGAGCCACAGGCACTGCCTTCCGCCGTAGTCCCCGCAACCTCATCTACTTTGG

Fig. 8 / continuation

AAACACAGTGTTACACATCCTCATCCTCCAGCCCAACAAAACCTTTGCCTGCCAGATGTACAACCTGTTGCTGTCCTACGACAGAC
ATGGGGACCACCTGCAGCCCCTGGACCTCGTGCCCAATCACCAGGGTCTCACCCCTTTCAAGCTGGCTGGAGTGGAGGGTAACACT
GTGATGTTTCAGCACCTGATGCAGAAGCGGAAGCACACCCAGTGGACGTATGGACCACTGACCTCGACTCTCTATGACCTCACAGA
GATCGACTCCTCAGGGGATGAGCAGTCCCTGCTGGAACTTATCATCACCACCAAGAAGCGGGAGGCTCGCCAGATCCTGGACCAGA
CGCCGGTGAAGGAGCTGGTGAGCCTCAAGTGGAAGCGGTACGGGCGGCCGTACTTCTGCATGCTGGGTGCCATATATCTGCTGTAC
ATCATCTGCTTCACCATGTGCTGCATCTACCGCCCCCTCAAGCCCAGGACCAATAACCGCACAAGCCCCCGGGACAACACCCTCTT
ACAGCAGAAGCTACTTCAGGAAGCCTACGTGACCCCTAAGGACGATATCCGGCTGGTCGGGGAGCTGGTGACTGTCATTGGGGCTA
TCATCATCCTGCTGGTAGAGGTTCCAGACATCTTCAGAATGGGGGTCACTCGCTTCTTTGGACAGACCATCCTTGGGGGCCCATTC
CATGTCCTCATCATCACCTATGCCTTCATGGTGCTGGTGACCATGGTGATGCGGCTCATCAGTGCCAGCGGGGAGGTGGTACCCAT
GTCCTTTGCACTCGTGCTGGGCTGGTGCAACGTCATGTACTTCGCCCGAGGATTCCAGATGCTAGGCCCCTTCACCATCATGATTC
AGAAGATGATTTTTGGCGACCTGATGCGATTCTGCTGGCTGATGGCTGTGGTCATCCTGGGCTTTGCTTCAGCCTTCTATATCATC
TTCCAGACAGAGGACCCCGAGGAGCTAGGCCACTTCTACGACTACCCCATGGCCCTGTTCAGCACCTTCGAGCTGGTCCTTACCAT
CATCGATGGCCCAGCCAACTACAACGTGGACCTGCCCTTCATGTACAGCATCACCTATGCTGCCTTTGCCATCATCGCCACACTGC
TCATGCTCAACCTCCTCATTGCCATGATGGGCGACACTCACTGGCGAGTGGCCCATGAGCGGGATGAGCTGTGGAGGGCCCAGATT
GTGGCCACCACGGTGATGCTGGAGCGGAAGCTGCCTCGCTGCCTGTGGCCTCGCTCCGGGATCTGCGGACGGGAGTATGGCCTGGG
GGACCGCTGGTTCCTGCGGGTGGAAGACAGGCAAGATCTCAACCGGCAGCGGATCCAACGCTACGCACAGGCCTTCCACACCCGGG
GCTCTGAGGATTTGGACAAAGACTCAGTGGAAAAACTAGAGCTGGGCTGTCCCTTCAGCCCCCACCTGTCCCTTCCTACGCCCTCA
GTGTCTCGAAGTACCTCCCGCAGCAGTGCCAATTGGGAAAGGCTTCGGCAAGGGACCCTGAGGAGAGACCTGCGTGGGATAATCAA
CAGGGGTCTGGAGGACGGGGAGAGCTGGGAATATCAGATCTGACTGCGTGTTCTCACTTCGCTTCCTGGAACTTGCTCTCATTTTC
CTGGGTGCATCAAACAAAACAAAAACCAAACACCCAGAGGTCTCATCTCCCAGGCCCCAGGGAGAAAGAGGAGTAGCATGAACGCC
AAGGAATGTACGTTGAGAATCACTGCTCCAGGCCTGCATTACTCCTTCAGCTCTGGGGCAGAGGAAGCCCAGCCCAAGCACGGGGC
TGGCAGGGCGTGAGGAACTCTCCTGTGGCCTGCTCATCACCCTTCCGACAGGAGCACTGCATGTCAGAGCACTTTAAAAACAGGCC
AGCCTGCTTGGGCCCTCGGTCTCCACCCCAGGGTCATAAGTGGGGAGAGAGCCCTTCCCAGGGCACCCAGGCAGGTGCAGGGAAGT
GCAGAGCTTGTGGAAAGCGTGTGAGTGAGGGAGACAGGAACGGCTCTGGGGGTGGGAAGTGGGGCTAGGTCTTGCCAACTCCATCT
TCAATAAAGTCGTTTTCGGATCCCTAAAAAAAAAAAAAAAAAAAAAAAAA

E.)

CACACATGGGGCCTCCCAGGAGTGCCCAGGACCTCGTGCTGTTGGCCTCTGAATCTATCGTCTCCAATCCGCTGTCCCACAGAAGC
CATATAACCCACCTCTCTGTAAATGCCAGGAGCCATGGGGGAAACAGCGCTACACATAGCAGCCCTCTATGACAACCTGGAGGCCG
CCATGGTGCTGATGGAGGCTGCCCCGGAGCTGGTCTTTGAGCCCATGACATCTGAGCTCTATGGAGGGTGAGGGCCCACGGGTCTG
GGGTGAAGAGCAGGAGTGACGTGGTTGGGTATTCAAGTCAGTCTCTGTGATGGATAATTTGGGAAAGACACAGGGGATCTGAGCCT
CCTACTCTTTTTSTCTTCTCTGTCTCCCTTCCGTGTCAGTCCCTGACTGCCCATCACTTGAACGCCTGCCCCCTGAAATGCCAGGG
GCCTAGAGAAGAGGAAGAGATGGGCAGCAGCTGGATCCCCTGGGAATCCTGAACACCCGAGAGCTCCCTGTTCTCCATCCCAGGCT
ACCCCTGAGGGAAAGAGACTAGGGGTGCATATGGGAGGGACCCCCTGCAGGATCCTAGGGGACAGACCCGTGACTGACAGCTGTCT
CTGGGCCAGGTCAGACTGCACTGCACATCGCTGTTGTGAACCAGAACATGAACCTGGTGCGAGCCCTGCTTGCCCGCAGGGCCAGT
GTCTCTGCCAGAGCCACAGGCACTGCCTTCCGCCGTAGTCCCTGCAACCTCATCTACTTTGGGGAGCACCCTTTGTCCTTTGCTGC
CTGTGTGAACAGTGAGGAGATCGTGCGGCTGCTCATTGAGCATGGAGCTGACATCCGGGCCCAGGACTCCCTGGATGTACAACCTG
TTGCTGTCCTACGACAGACATGGGGACCACCTGCAGCCCCTGGACCTCGTGCCCAATCACCAGGGTCTCACCCCTTTCAAGCTGGC
TGGAGTGGAGGGTAACACTGTGATGTTTCAGCACCTGATGCAGAAGCGGAAGCACACCCAGTGGACGTATGGACCACTGACCTCGA
CTCTCTATGACCTCACAGAGATCGACTCCTCAGGGGATGAGCAGTCCCTGCTGGAACTTATCATCACCACCAAGAAGCGGGAGGCT
CGCCAGATCCTGGACCAGACGCCGGTGAAGGAGCTGGTGAGCCTCAAGTGGAAGCGGTACGGGCGGCCGTACTTCTGCATGCTGGG
TGCCATATATCTGCTGTACATCATCTGCTTCACCATGTGCTGCATCTACCGCCCCCTCAAGCCCAGGACCAATAACCGCACGAGCC
CCCGGGACAACACCCTCTTACAGCAGAAGCTACTTCAGGAAGCCTACATGACCCCTAAGGACGATATCCGGCTGGTCGGGGAGCTG
GTGACTGTCATTGGGGCTATCATCATCCTGCTGGTAGAGGTTCCAGACATCTTCAGAATGGGGGTCACTCGCTTCTTTGGACAGAC
CATCCTTGGGGGCCCATTCCATGTCCTCATCATCACCTATGCCTTCATGGTGCTGGTGACCATGGTGATGCGGCTCATCAGTGCCA
GCGGGGAGGTGGTACCCATGTCCTTTGCACTCGTGCTGGGCTGGTGCAACGTCATGTACTTCGCCCGAGGATTCCAGATGCTAGGC
CCCTTCACCATCATGATTCAGAAGATGATTTTTGGCGACCTGATGCGATTCTGCTGGCTGATGGCTGTGGTCATCCTGGGCTTTGC
TTCAGCCTTCTATATCATCTTCCAGACAGAGGACCCCGAGGAGCTAGGCCACTTCTACGACTACCCCATGGCCCTGTTCAGCACCT
TCGAGCTGGTCCTTACCATCATCGATGGCCCAGCCAACTACAACGTGGACCTGCCCTTCATGTACAGCATCACCTATGCTGCCTTT
GCCATCATCGCCACACTGCTCATGCTCAACCTCCTCATTGCCATGATGGGCGACACTCACTGGCGAGTGGCCCATGAGCGGGATGA
GCTGTGGAGGGCCCAGATTGTGGCCACCACGGTGATGCTGGAGCGGAAGCTGCCTCGCTGCCTGTGGCCTCGCTCCGGGATCTGCG
GACGGGAGTATGGCCTGGGAGACCGCTGGTTCCTGCGGGTGGAAGACAGGCAAGATCTCAACCGGCAGCGGATCCAACGCTACGCA
CAGGCCTTCCACACCCGGGGCTCTGAGGATTTGGACAAAGACTCAGTGGAAAAACTAGAGCTGGGCTGTCCCTTCAGCCCCCACCT
GTCCCTTCCTATGCCCTCAGTGTCTCGAAGTACCTCCCGCAGCAGTGCCAATTGGGAAAGGCTTCGGCAAGGGACCCTGAGGAGAG
ACCTGCGTGGGATAATCAACAGGGGTCTGGAGGACGGGGAGAGCTGGGAATATCAGATCTGACTGCGTGTTCTCACTTCGCTTCCT
GGAACTTGCTCTCATTTTCCTGGGTGCATCAAACAAAACAAAAACCAAACACCCAGAGGTCTCATCTCCCAGGCCCCAGGGAGAAA
GAGGAGTAGCATGAACGCCAAGGAATGTACGTTGAGAATCACTGCTCCAGGCCTGCATTACTCCTTCAGCTCTGGGGCAGAGGAAG
CCCAGCCCAAGCACGGGGCTGGCAGGGCGTGAGGAACTCTCCTGTGGCCTGCTCATCACCCTTCCGACAGGAGCACTGCATGTCAG
AGCACTTTAAAAACAGGCCAGCCTGCTTGGGCCCTCGGTCTCCACCCCAGGGTCATAAGTGGGGAGAGAGCCCTTCCCAGGGCACC

Fig. 8 / continuation 5

CAGGCAGGTGCAGGGAAGTGCAGAGCTTGTGGAAAGCGTGTGAGTGAGGGAGACAGGAACGGCTCTGGGGGTGGGAAGTGGGGCTA
GGTCTTGCCAACTCCATCTTCAATAAAGTCGTTTTCGGATCCCTAAAAAAAAAAAAAAAAAAAAAAAAAAA

Figure 9:

A.

```
       10                  30                  50
CGGGGCCCTGGGCTGCAGGAGGTTGCGGCGGCCGCGGCAGCATGGTGGTGCCGGAGAAGG
                                          M  V  V  P  E  K  E
       70                  90                 110
AGCAGAGCTGGATCCCCAAGATCTTCAAGAAGAAGACCTGCACGACGTTCATAGTTGACT
  Q  S  W  I  P  K  I  F  K  K  K  T  C  T  T  F  I  V  D  S
      130                 150                 170
CCACAGATCCGGGAGGGACCTTGTGCCAGTGTGGGCGCCCCCGGACCGCCCACCCCGCAG
  T  D  P  G  G  T  L  C  Q  C  G  R  P  R  T  A  H  P  A  V
      190                 210                 230
TGGCCATGGAGGATGCCTTCGGGGCAGCCGTGGTGACCGTGTGGGACAGCGATGCACACA
  A  M  E  D  A  F  G  A  A  V  V  T  V  W  D  S  D  A  H  T
      250                 270                 290
CCACGGAGAAGCCCACCGATGCCTACGGAGAGCTGGACTTCACGGGGGCCGGCCGCAAGC
  T  E  K  P  T  D  A  Y  G  E  L  D  F  T  G  A  G  R  K  H
      310                 330                 350
ACAGCAATTTCCTCCGGCTCTCTGACCGAACGGATCCAGCTGCAGTTTATAGTCTGGTCA
  S  N  F  L  R  L  S  D  R  T  D  P  A  A  V  Y  S  L  V  T
      370                 390                 410
CACGCACATGGGGCTTCCGTGCCCCGAACCTGGTGGTGTCAGTGCTGGGGGGATCGGGGG
  R  T  W  G  F  R  A  P  N  L  V  V  S  V  L  G  G  S  G  G
      430                 450                 470
GCCCCGTCCTCCAGACCTGGCTGCAGGACCTGCTGCGTCGTGGGCTGGTGCGGGCTGCCC
  P  V  L  Q  T  W  L  Q  D  L  L  R  R  G  L  V  R  A  A  Q
      490                 510                 530
AGAGCACAGGAGCCTGGATTGTCACTGGGGGTCTGCACACGGGCATCGGCCGGCATGTTG
  S  T  G  A  W  I  V  T  G  G  L  H  T  G  I  G  R  H  V  G
      550                 570                 590
GTGTGGCTGTACGGGACCATCAGATGGCCAGCACTGGGGGCACCAAGGTGGTGGCCATGG
  V  A  V  R  D  H  Q  M  A  S  T  G  G  T  K  V  V  A  M  G
      610                 630                 650
GTGTGGCCCCCTGGGGTGTGGTCCGGAATAGAGACACCCTCATCAACCCCAAGGGCTCGT
  V  A  P  W  G  V  V  R  N  R  D  T  L  I  N  P  K  G  S  F
      670                 690                 710
TCCCTGCGAGGTACCGGTGGCGCGGTGACCCGGAGGACGGGGTCCAGTTTCCCCTGGACT
  P  A  R  Y  R  W  R  G  D  P  E  D  G  V  Q  F  P  L  D  Y
      730                 750                 770
ACAACTACTCGGCCTTCTTCCTGGTGGACGACGGCACACACGGCTGCCTGGGGGGCGAGA
  N  Y  S  A  F  F  L  V  D  D  G  T  H  G  C  L  G  G  E  N
      790                 810                 830
ACCGCTTCCGCTTGCGCCTGGAGTCCTACATCTCACAGCAGAAGACGGGCGTGGGAGGGA
  R  F  R  L  R  L  E  S  Y  I  S  Q  Q  K  T  G  V  G  G  T
      850                 870                 890
CTGGAATTGACATCCCTGTCCTGCTCCTCCTGATTGATGGTGATGAGAAGATGTTGACGC
  G  I  D  I  P  V  L  L  L  L  I  D  G  D  E  K  M  L  T  R
      910                 930                 950
GAATAGAGAACGCCACCCAGGCTCAGCTCCCATGTCTCCTCGTGGCTGGCTCAGGGGGAG
  I  E  N  A  T  Q  A  Q  L  P  C  L  L  V  A  G  S  G  G  A
      970                 990                1010
CTGCGGACTGCCTGGCGGAGACCCTGGAAGACACTCTGGCCCCAGGGAGTGGGGGAGCCA
  A  D  C  L  A  E  T  L  E  D  T  L  A  P  G  S  G  G  A  R
     1030                1050                1070
GGCAAGGCGAAGCCCGAGATCGAATCAGGCGTTTCTTTCCCAAAGGGGACCTTGAGGTCC
```

Fig. 9 / continuation 1

```
  Q  G  E  A  R  D  R  I  R  R  F  F  P  K  G  D  L  E  V  L
     1090              1110              1130
TGCAGGCCCAGGTGGAGAGGATTATGACCCGGAAGGAGCTCCTGACAGTCTATTCTTCTG
  Q  A  Q  V  E  R  I  M  T  R  K  E  L  L  T  V  Y  S  S  E
     1150              1170              1190
AGGATGGGTCTGAGGAATTCGAGACCATAGTTTTGAAGGCCCTTGTGAAGGCCTGTGGGA
  D  G  S  E  E  F  E  T  I  V  L  K  A  L  V  K  A  C  G  S
     1210              1230              1250
GCTCGGAGGCCTCAGCCTACCTGGATGAGCTGCGTTTGGCTGTGGCTTGGAACCGCGTGG
  S  E  A  S  A  Y  L  D  E  L  R  L  A  V  A  W  N  R  V  D
     1270              1290              1310
ACATTGCCCAGAGTGAACTCTTTCGGGGGGACATCCAATGGCGGTCCTTCCATCTCGAAG
  I  A  Q  S  E  L  F  R  G  D  I  Q  W  R  S  F  H  L  E  A
     1330              1350              1370
CTTCCCTCATGGACGCCCTGCTGAATGACCGGCCTGAGTTCGTGCGCTTGCTCATTTCCC
  S  L  M  D  A  L  L  N  D  R  P  E  F  V  R  L  L  I  S  H
     1390              1410              1430
ACGGCCTCAGCCTGGGCCACTTCCTGACCCCGATGCGCCTGGCCCAACTCTACAGCGCGG
  G  L  S  L  G  H  F  L  T  P  N  R  L  A  Q  L  Y  S  A  A
     1450              1470              1490
CGCCCTCCAACTCGCTCATCCGCAACCTTTTGGACCAGGCGTCCCACAGCGCAGGCACCA
  P  S  N  S  L  I  R  N  L  L  D  Q  A  S  H  S  A  G  T  K
     1510              1530              1550
AAGCCCCAGCCCTAAAAGGGGGAGCTGCGGAGCTCCGGCCCCCTGACGTGGGGCATGTGC
  A  P  A  L  K  G  G  A  A  E  L  R  P  P  D  V  G  H  V  L
     1570              1590              1610
TGAGGATGCTGCTGGGGAAGATGTGCGCGCCGAGGTACCCCTCCGGGGGCGCCTGGGACC
  R  M  L  L  G  K  M  C  A  P  R  Y  P  S  G  G  A  W  D  P
     1630              1650              1670
CTCACCCAGGCCAGGGCTTCGGGGAGAGCATGTATCTGCTCTCGGACAAGGCCACCTCGC
  H  P  G  Q  G  F  G  E  S  M  Y  L  L  S  D  K  A  T  S  P
     1690              1710              1730
CGCTCTCGCTGGATGCTGGCCTCGGGCAGGCCCCCTGGAGCGACCTGCTTCTTTGGGCAC
  L  S  L  D  A  G  L  G  Q  A  P  W  S  D  L  L  L  W  A  L
     1750              1770              1790
TGTTGCTGAACAGGGCACAGATGGCCATGTACTTCTGGGAGATGGGTTCCAATGCAGTTT
  L  L  N  R  A  Q  M  A  M  Y  F  W  E  M  G  S  N  A  V  S
     1810              1830              1850
CCTCAGCTCTTGGGGCCTGTTTGCTGCTCCGGGTGATGGCACGCCTGGAGCCTGACGCTG
  S  A  L  G  A  C  L  L  L  R  V  M  A  R  L  E  P  D  A  E
     1870              1890              1910
AGGAGGCAGCACGGAGGAAAGACCTGGCGTTCAAGTTTGAGGGGATGGGCGTTGACCTCT
  E  A  A  R  R  K  D  L  A  F  K  F  E  G  M  G  V  D  L  F
     1930              1950              1970
TTGGCGAGTGCTATCGCAGCAGTGAGGTGAGGGCTGCCCGCCTCCTCCTCCGTCGCTGCC
  G  E  C  Y  R  S  S  E  V  R  A  A  R  L  L  L  R  R  C  P
     1990              2010              2030
CGCTCTGGGGGGATGCCACTTGCCTCCAGCTGGCCATGCAAGCTGACGCCCGTGCCTTCT
  L  W  G  D  A  T  C  L  Q  L  A  M  Q  A  D  A  R  A  F  F
     2050              2070              2090
TTGCCCAGGATGGGGTACAGTCTCTGCTGACACAGAAGTGGTGGGGAGATATGGCCAGCA
  A  Q  D  G  V  Q  S  L  L  T  Q  K  W  W  G  D  M  A  S  T
     2110              2130              2150
CTACACCCATCTGGGCCCTGGTTCTCGCCTTCTTTTGCCCTCCACTCATCTACACCCGCC
  T  P  I  W  A  L  V  L  A  F  F  C  P  P  L  I  Y  T  R  L
     2170              2190              2210
TCATCACCTTCAGGAAATCAGAAGAGGAGCCCACACGGGAGGAGCTAGAGTTTGACATGG
  I  T  F  R  K  S  E  E  E  P  T  R  E  E  L  E  F  D  M  D
     2230              2250              2270
ATAGTGTCATTAATGGGGAAGGGCCTGTCGGGACGGCGGACCCAGCCGAGAAGACGCCGC
  S  V  I  N  G  E  G  P  V  G  T  A  D  P  A  E  K  T  P  L
     2290              2310              2330
```

Fig. 9 / continuation 2

```
TGGGGGTCCCGCGCCAGTCGGGCCGTCCGGGTTGCTGCGGGGGCCGCTGCGGGGGGCGCC
  G  V  P  R  Q  S  G  R  P  G  C  C  G  G  R  C  G  G  R  R
     2350                2370                2390
GGTGCCTACGCCGCTGGTTCCACTTCTGGGGCGTGCCGGTGACCATCTTCATGGGCAACG
  C  L  R  R  W  F  H  F  W  G  V  P  V  T  I  F  M  G  N  V
     2410                2430                2450
TGGTCAGCTACCTGCTGTTCCTGCTGCTTTTCTCGCGGGTGCTGCTCGTGGATTTCCAGC
  V  S  Y  L  L  F  L  L  L  F  S  R  V  L  L  V  D  F  Q  P
     2470                2490                2510
CGGCGCCGCCCGGCTCCCTGGAGCTGCTGCTCTATTTCTGGGCTTTCACGCTGCTGTGCG
  A  P  P  G  S  L  E  L  L  L  Y  F  W  A  F  T  L  L  C  E
     2530                2550                2570
AGGAACTGCGCCAGGGCCTGAGCGGAGGCGGGGGCAGCCTCGCCAGCGGGGGCCCCGGGC
  E  L  R  Q  G  L  S  G  G  G  G  S  L  A  S  G  G  P  G  P
     2590                2610                2630
CTGGCCATGCCTCACTGAGCCAGCGCCTGCGCCTCTACCTCGCCGACAGCTGGAACCAGT
  G  H  A  S  L  S  Q  R  L  R  L  Y  L  A  D  S  W  N  Q  C
     2650                2670                2690
GCGACCTAGTGGCTCTCACCTGCTTCCTCCTGGGCGTGGGCTGCCGGCTGACCCCGGGTT
  D  L  V  A  L  T  C  F  L  L  G  V  G  C  R  L  T  P  G  L
     2710                2730                2750
TGTACCACCTGGGCCGCACTGTCCTCTGCATCGACTTCATGGTTTTCACGGTGCGGCTGC
  Y  H  L  G  R  T  V  L  C  I  D  F  M  V  F  T  V  R  L  L
     2770                2790                2810
TTCACATCTTCACGGTCAACAAACAGCTGGGGCCCAAGATCGTCATCGTGAGCAAGATGA
  H  I  F  T  V  N  K  Q  L  G  P  K  I  V  I  V  S  K  M  M
     2830                2850                2870
TGAAGGACGTGTTCTTCTTCCTCTTCTTCCTCGGCGTGTGGCTGGTAGCCTATGGCGTGG
  K  D  V  F  F  F  L  F  F  L  G  V  W  L  V  A  Y  G  V  A
     2890                2910                2930
CCACGGAGGGGCTCCTGAGGCCACGGGACAGTGACTTCCCAAGTATCCTGCGCCGCGTCT
  T  E  G  L  L  R  P  R  D  S  D  F  P  S  I  L  R  R  V  F
     2950                2970                2990
TCTACCGTCCCTACCTGCAGATCTTCGGGCAGATTCCCCAGGAGGACATGGACGTGGCCC
  Y  R  P  Y  L  Q  I  F  G  Q  I  P  Q  E  D  M  D  V  A  L
     3010                3030                3050
TCATGGAGCACAGCAACTGCTCGTCGGAGCCCGGCTTCTGGGCACACCCTCCTGGGGCCC
  M  E  H  S  N  C  S  S  E  P  G  F  W  A  H  P  P  G  A  Q
     3070                3090                3110
AGGCGGGCACCTGCGTCTCCCAGTATGCCAACTGGCTGGTGGTGCTGCCTCCTCGTCATCT
  A  G  T  C  V  S  Q  Y  A  N  W  L  V  V  L  L  L  V  I  F
     3130                3150                3170
TCCTGCTCGTGGCCAACATCCTGCTGGTCAACTTGCTCATTGCCATGTTCAGTTACACAT
  L  L  V  A  N  I  L  L  V  N  L  L  I  A  M  F  S  Y  T  F
     3190                3210                3230
TCGGCAAAGTACAGGGCAACAGCGATCTCTACTGGAAGGCGCAGCGTTACCGCCTCATCC
  G  K  V  Q  G  N  S  D  L  Y  W  K  A  Q  R  Y  R  L  I  R
     3250                3270                3290
GGGAATTCCACTCTCGGCCCGCGCTGGCCCCGCCCTTTATCGTCATCTCCCACTTGCGCC
  E  F  H  S  R  P  A  L  A  P  P  F  I  V  I  S  H  L  R  L
     3310                3330                3350
TCCTGCTCAGGCAATTGTGCAGGCGACCCCGGAGCCCCCAGCCGTCCTCCCCGGCCCTCG
  L  L  R  Q  L  C  R  R  P  R  S  P  Q  P  S  S  P  A  L  E
     3370                3390                3410
AGCATTTCCGGGTTTACCTTTCTAAGGAAGCCGAGCGGAAGCTGCTAACGTGGGAATCGG
  H  F  R  V  Y  L  S  K  E  A  E  R  K  L  L  T  W  E  S  V
     3430                3450                3470
TGCATAAGGAGAACTTTCTGCTGGCACGCGCTAGGGACAAGCGGGAGAGCGACTCCGAGC
  H  K  E  N  F  L  L  A  R  A  R  D  K  R  E  S  D  S  E  R
     3490                3510                3530
GTCTGAAGCGCACGTCCCAGAAGGTGGACTTGGCACTGAAACAGCTGGGACACATCCGCG
  L  K  R  T  S  Q  K  V  D  L  A  L  K  Q  L  G  H  I  R  E
```

Fig. 9 / continuation 3

```
     3550                3570                3590
AGTACGAACAGCGCCTGAAAGTGCTGGAGCGGGAGGTCCAGCAGTGTAGCCGCGTCCTGG
  Y  E  Q  R  L  K  V  L  E  R  E  V  Q  Q  C  S  R  V  L  G
     3610                3630                3650
GGTGGGTGGCCGAGGCCCTGAGCCGCTCTGCCTTGCTGCCCCCAGGTGGGCCGCCACCCC
  W  V  A  E  A  L  S  R  S  A  L  L  P  P  G  G  P  P  P  P
     3670                3690                3710
CTGACCTGCCTGGGTCCAAAGACTGAGCCCTGCTGGCGGACTTCAAGGAGAAGCCCCCAC
  D  L  P  G  S  K  D  *
     3730                3750                3770
AGGGGATTTTGCTCCTAGAGTAAGGCTCATCTGGGCCTCGGCCCCCGCACCTGGTGGCCT
     3790                3810                3830
TGTCCTTGAGGTGAGCCCCATGTCCATCTGGGCCACTGTCAGGACCACCTTTGGGAGTGT
     3850                3870                3890
CATCCTTACAAACCACAGCATGCCCGGCTCCTCCCAGAACCAGTCCCAGCCTGGGAGGAT
     3910                3930                3950
CAAGGCCTGGATCCCGGGCCGTTATCCATCTGGAGGCTGCAGGGTCCTTGGGGTAACAGG
     3970                3990                4010
GACCACAGACCCCTCACCACTCACAGATTCCTCACACTGGGGAAATAAAGCCATTTCAGA
     4030
GGAAAAAAAAAAAAAAAAAAAA
```

```
MVVPEKEQSWIPKIFKKKTCTTFIVDSTDPGGTLCQCGRPRTAHPAVAMEDAFGAAVVTVWDSDAHTTEKPTDAYELDFTGAG
SNFLRLSDRTDPAAVYSLVTRTWGFRAPNLVVSVLGGSGGPVLQTWLQDLLRRGLVRAAQSTGAWIVTGGLHTGIGRHVGVAV
QMASTGGTKVVAMGVAPWGVVRNRDTLINPKGSFPARYRWRGDPEDGVQFPLDYNYSAFLVDDGTHGCLGGGENRFRLRLESY
QKTGVGGTGIDIPVLLLLIDGDEKNLTRIENATQAHVPCLLVAGSRGLGMPGGTLEAHLAQDGDHKANQSTNQLLLPKDLSLQ
SIDRKTLQSYSERLAVAWNRVDIAQSELFRGDIQWRSFHLEASLMDALLNDRPEFVRLLISHGLSLGHFLTPMRLAQLYSAAE
LIRNLLDQASHSAGTKAPALKGGAAELRPPDVGHVLRMLLGKMCAPRYPSGGAWDPHPGQGFGESMYLLSDKATSPLSLDAGI
PWSDLLLWALLLNRAQMAMYFWEMGSNAVSSALGACLLLRVMARLEPDAEEAARRKDLAFKFEGMGVDLFGECYRSSEVRAAF
RRCPLWGDATCLQLAMQADARAFFAQDGVQSLLTQKWWGDMASTTPIWALVLAFFCPPLIYTRLITFRKSEEEPTREELEFDM
INGEGPVGTADPAEKTPLGVPRQSGRPGCCGGRCGGRRCLRRWFHFWGVPVTIFMGNVVSYLLFLLLFSRVLLVDFQPAPPGS
LLYFWAFTLLCEELRQGLSGGGGSLASGGPGPGHASLSQRLRLYLADSWNQCDLVALTCFLLGVGCRLTPGLYHLGRTVLCII
FTVRLLHIFTVNKQLGPKIVIVSKMMKDVFFFLFFLGVWLVAYGVATEGLLRPRDSDFPSILRRVFYRPYLQIFGQIPQEDMI
MEHSNCSSEPGFWAHPPGAQAGTCVSQYANWLVVLLLVIFLLVANILLVNLLIAMFSYTFGKVQGNSDLYWKAQRYRLIREFH
ALAPPFIVISHLRLLLRQLCRRPRSPQPSSPALEHFRVYLSKEAERKLLTWESVHKENFLLARARDKRESDSERLKRTSQKVI
KQLGHIREYEQRLKVLEREVQQCSRVLGWVAEALSRSALLPPGGPPPPDLPGSKD
```

B.)

```
        10                  30                  50
ATCCAATGGCGGTCCTTCCATCTCGAAGCTTCCCTCATGGACGCCCTGCTGAATGACCGG
        70                  90                  110
CCTGAGTTCGTGCGCTTGCTCATTTCCCACGGCCTCAGCCTGGGCCACTTCCTGACCCCG
       130                 150                 170
ATGCGCCTGGCCCAACTCTACAGCGCGGCGCCCTCCAACTCGCTCATCCGCAACCTTTTG
       190                 210                 230
GACCAGGCGTCCCACAGCGCAGGCACCAAAGCCCCAGCCCTAAAAGGGGGAGCTGCGGAG
       250                 270                 290
CTCCGGCCCCCTGACGTGGGGCATGTGCTGAGGATGCTGCTGGGGAAGATGTGCGCGCCG
       310                 330                 350
AGATGTATCTGCTCTCGGACAAGGCCACCTCGCCGCTCTCGCTGGATGCTGGCCTCGGGC
  M  Y  L  L  S  D  K  A  T  S  P  L  S  L  D  A  G  L  G  Q
       370                 390                 410
AGGCCCCCTGGAGCGACCTGCTTCTTTGGGCACTGTTGCTGAACAGGGCACAGATGGCCA
  A  P  W  S  D  L  L  L  W  A  L  L  L  N  R  A  Q  M  A  M
       430                 450                 470
TGTACTTCTGGGAGATGGGTTCCAATGCAGTTTCCTCAGCTCTTGGGGCCTGTTTGCTGC
  Y  F  W  E  M  G  S  N  A  V  S  S  A  L  G  A  C  L  L  L
```

Fig. 9 / continuation 4

```
      490                 510                 530
TCCGGGTGATGGCACGCCTGGAGCCTGACGCTGAGGAGGCAGCACGGAGGAAAGACCTGG
  R  V  M  A  R  L  E  P  D  A  E  E  A  A  R  R  K  D  L  A
      550                 570                 590
CGTTCAAGTTTGAGGGGATGGGCGTTGACCTCTTTGGCGAGTGCTATCGCAGCAGTGAGG
  F  K  F  E  G  M  G  V  D  L  F  G  E  C  Y  R  S  S  E  V
      610                 630                 650
TGAGGGCTGCCCGCCTCCTCCTCCGTCGCTGCCCGCTCTGGGGGGATGCCACTTGCCTCC
  R  A  A  R  L  L  L  R  R  C  P  L  W  G  D  A  T  C  L  Q
      670                 690                 710
AGCTGGCCATGCAAGCTGACGCCCGTGCCTTCTTTGCCCAGGATGGGGTACAGTCTCTGC
  L  A  M  Q  A  D  A  R  A  F  F  A  Q  D  G  V  Q  S  L  L
      730                 750                 770
TGACACAGAAGTGGTGGGGAGATATGGCCAGCACTACACCCATCTGGGCCCTGGTTCTCG
  T  Q  K  W  W  G  D  M  A  S  T  T  P  I  W  A  L  V  L  A
      790                 810                 830
CCTTCTTTTGCCCTCCACTCATCTACACCCGCCTCATCACCTTCAGGAAATCAGAAGAGG
  F  F  C  P  P  L  I  Y  T  R  L  I  T  F  R  K  S  E  E  E
      850                 870                 890
AGCCCACACGGGAGGAGCTAGAGTTTGACATGGATAGTGTCATTAATGGGGAAGGGCCTG
  P  T  R  E  E  L  E  F  D  M  D  S  V  I  N  G  E  G  P  V
      910                 930                 950
TCGGGACGGCGGACCCAGCCGAGAAGACGCCGCTGGGGGTCCCGCGCCAGTCGGGCCGTC
  G  T  A  D  P  A  E  K  T  P  L  G  V  P  R  Q  S  G  R  P
      970                 990                1010
CGGGTTGCTGCGGGGGCCGCTGCGGGGGGCGCCGGTGCCTACGCCGCTGGTTCCACTTCT
  G  C  C  G  G  R  C  G  G  R  R  C  L  R  R  W  F  H  F  W
     1030                1050                1070
GGGGCGTGCCGGTGACCATCTTCATGGGCAACGTGGTCAGCTACCTGCTGTTCCTGCTGC
  G  V  P  V  T  I  F  M  G  N  V  V  S  Y  L  L  F  L  L  L
     1090                1110                1130
TTTTCTCGCGGGTGCTGCTCGTGGATTTCCAGCCGGCGCCGCCCGGCTCCCTGGAGCTGC
  F  S  R  V  L  L  V  D  F  Q  P  A  P  P  G  S  L  E  L  L
     1150                1170                1190
TGCTCTATTTCTGGGCTTTCACGCTGCTGTGCGAGGAACTGCGCCAGGGCCTGAGCGGAG
  L  Y  F  W  A  F  T  L  L  C  E  E  L  R  Q  G  L  S  G  G
     1210                1230                1250
GCGGGGGCAGCCTCGCCAGCGGGGGCCCCGGGCCTGGCCATGCCTCACTGAGCCAGCGCC
  G  G  S  L  A  S  G  G  P  G  P  G  H  A  S  L  S  Q  R  L
     1270                1290                1310
TGCGCCTCTACCTCGCCGACAGCTGGAACCAGTGCGACCTAGTGGCTCTCACCTGCTTCC
  R  L  Y  L  A  D  S  W  N  Q  C  D  L  V  A  L  T  C  F  L
     1330                1350                1370
TCCTGGGCGTGGGCTGCCGGCTGACCCCGGGTTTGTACCACCTGGGCCGCACTGTCCTCT
  L  G  V  G  C  R  L  T  P  G  L  Y  H  L  G  R  T  V  L  C
     1390                1410                1430
GCATCGACTTCATGGTTTTCACGGTGCGGCTGCTTCACATCTTCACGGTCAACAAACAGC
  I  D  F  M  V  F  T  V  R  L  L  H  I  F  T  V  N  K  Q  L
     1450                1470                1490
TGGGGCCCAAGATCGTCATCGTGAGCAAGATGATGAAGGACGTGTTCTTCTTCCTCTTCT
  G  P  K  I  V  I  V  S  K  M  M  K  D  V  F  F  F  L  F  F
     1510                1530                1550
TCCTCGGCGTGTGGCTGGTAGCCTATGGCGTGGCCACGGAGGGGCTCCTGAGGCCACGGG
  L  G  V  W  L  V  A  Y  G  V  A  T  E  G  L  L  R  P  R  D
     1570                1590                1610
ACAGTGACTTCCCAAGTATCCTGCGCCGCGTCTTCTACCGTCCCTACCTGCAGATCTTCG
  S  D  F  P  S  I  L  R  R  V  F  Y  R  P  Y  L  Q  I  F  G
     1630                1650                1670
GGCAGATTCCCCAGGAGGACATGGACGTGGCCCTCATGGAGCACAGCAACTGCTCGTCGG
  Q  I  P  Q  E  D  M  D  V  A  L  M  E  H  S  N  C  S  S  E
     1690                1710                1730
AGCCCGGCTTCTGGGCACACCCTCCTGGGGCCCAGGCGGGCACCTGCGTCTCCCAGTATG
```

Fig. 9 / continuation 5

```
  P  G  F  W  A  H  P  P  G  A  Q  A  G  T  C  V  S  Q  Y  A
     1750                 1770                 1790
CCAACTGGCTGGTGGTGCTGCTCCTCGTCATCTTCCTGCTCGTGGCCAACATCCTGCTGG
  N  W  L  V  V  L  L  L  V  I  F  L  L  V  A  N  I  L  L  V
     1810                 1830                 1850
TCAACTTGCTCATTGCCATGTTCAGTTACACATTCGGCAAAGTACAGGGCAACAGCGATC
  N  L  L  I  A  M  F  S  Y  T  F  G  K  V  Q  G  N  S  D  L
     1870                 1890                 1910
TCTACTGGAAGGCGCAGCGTTACCGCCTCATCCGGGAATTCCACTCTCGGCCCGCGCTGG
  Y  W  K  A  Q  R  Y  R  L  I  R  E  F  H  S  R  P  A  L  A
     1930                 1950                 1970
CCCCGCCCTTTATCGTCATCTCCCACTTGCGCCTCCTGCTCAGGCAATTGTGCAGGCGAC
  P  P  F  I  V  I  S  H  L  R  L  L  L  R  Q  L  C  R  R  P
     1990                 2010                 2030
CCCGGAGCCCCCAGCCGTCCTCCCCGGCCCTCGAGCATTTCCGGGTTTACCTTTCTAAGG
  R  S  P  Q  P  S  S  P  A  L  E  H  F  R  V  Y  L  S  K  E
     2050                 2070                 2090
AAGCCGAGCGGAAGCTGCTAACGTGGGAATCGGTGCATAAGGAGAACTTTCTGCTGGCAC
  A  E  R  K  L  L  T  W  E  S  V  H  K  E  N  F  L  L  A  R
     2110                 2130                 2150
GCGCTAGGGACAAGCGGGAGAGCGACTCCGAGCGTCTGAAGCGCACGTCCCAGAAGGTGG
  A  R  D  K  R  E  S  D  S  E  R  L  K  R  T  S  Q  K  V  D
     2170                 2190                 2210
ACTTGGCACTGAAACAGCTGGGACACATCCGCGAGTACGAACAGCGCCTGAAAGTGCTGG
  L  A  L  K  Q  L  G  H  I  R  E  Y  E  Q  R  L  K  V  L  E
     2230                 2250                 2270
AGCGGGAGGTCCAGCAGTGTAGCCGCGTCCTGGGGTGGGTGGCCGAGGCCCTGAGCCGCT
  R  E  V  Q  Q  C  S  R  V  L  G  W  V  A  E  A  L  S  R  S
     2290                 2310                 2330
CTGCCTTGCTGCCCCCAGGTGGGCCGCCACCCCCTGACCTGCCTGGGTCCAAAGACTGAG
  A  L  L  P  P  G  G  P  P  P  P  D  L  P  G  S  K  D  *
     2350                 2370                 2390
CCCTGCTGGCGGACTTCAAGGAGAAGCCCCCACAGGGGATTTTGCTCCTAGAGTAAGGCT
     2410                 2430                 2450
CATCTGGGCCTCGGCCCCCGCACCTGGTGGCCTTGTCCTTGAGGTGAGCCCCATGTCCAT
     2470                 2490                 2510
CTGGGCCACTGTCAGGACCACCTTTGGGAGTGTCATCCTTACAAACCACAGCATGCCCGG
     2530                 2550                 2570
CTCCTCCCAGAACCAGTCCCAGCCTGGGAGGATCAAGGCCTGGATCCCGGGCCGTTATCC
     2590                 2610                 2630
ATCTGGAGGCTGCAGGGTCCTTGGGGTAACAGGGACCACAGACCCCTCACCACTCACAGA
     2650                 2670                 2690
TTCCTCACACTGGGGAAATAAAGCCATTTCAGAGGAAAAAAAAAAAAAAAAAAA
```

```
MYLLSDKATSPLSLDAGLGQAPWSDLLLWALLLNRAQMAMYFWEMGSNAVSSALGACLLLRVMARLEPDAEEAARRKDLAFKFEGM
GVDLFGECYRSSEVRAARLLLRRCPLWGDATCLQLAMQADARAFFAQDGVQSLLTQKWWGDMASTTPIWALVLAFFCPPLIYTRLI
TFRKSEEEPTREELEFDMDSVINGEGPVGTADPAEKTPLGVPRQSGRPGCCGGRCGGRRCLRRWFHFWGVPVTIFMGNVVSYLLFL
LLFSRVLLVDFQPAPPGSLELLLYFWAFTLLCEELRQGLSGGGGSLASGGPGPGHASLSQRLRLYLADSWNQCDLVALTCFLLGVG
CRLTPGLYHLGRTVLCIDFMVFTVRLLHIFTVNKQLGPKIVIVSKMMKDVFFFLFFLGVWLVAYGVATEGLLRPRDSDFPSILRRV
FYRPYLQIFGQIPQEDMDVALMEHSNCSSEPGFWAHPPGAQAGTCVSQYANWLVVLLLVIFLLVANILLVNLLIAMFSYTFGKVQG
NSDLYWKAQRYRLIREFHSRPALAPPFIVISHLRLLLRQLCRRPRSPQPSSPALEHFRVYLSKEAERKLLTWESVHKENFLLARAR
DKRESDSERLKRTSQKVDLALKQLGHIREYEQRLKVLEREVQQCSRVLGWVAEALSRSALLPPGGPPPPDLPGSKD
```

Fig. 10

A)

```
        10                  30                  50
ATTAAAGTTTATAAAACAGTGGCTGGATGGTTGGAGGATGCAGGTGGACAGAAGACGTGG
                          M  V  G  G  C  R  W  T  E  D  V  E
        70                  90                  110
AGCCTGCAGAAGTAAAGGAAAAGATGTCCTTTCGGGCAGCCAGGCTCAGCATGAGGAACA
  P  A  E  V  K  E  K  M  S  F  R  A  A  R  L  S  M  R  N  R
      130                 150                 170
GAAGGAATGACACTCTGGACAGCACCCGGACCCTGTACTCCAGCGCGTCTCGGAGCACAG
  R  N  D  T  L  D  S  T  R  T  L  Y  S  S  A  S  R  S  T  D
      190                 210                 230
ACTTGTCTTACAGTGAAAGCGCCAGCTTCTACGCTGCCTTCAGGACACAGACGTGCCCAA
  L  S  Y  S  E  S  A  S  F  Y  A  A  F  R  T  Q  T  C  P  I
      250                 270                 290
TCATGGCTTCTTGGGACTTGGTGAATTTTATTCAAGCAAATTTTAAGAAACGAGAATGTG
  M  A  S  W  D  L  V  N  F  I  Q  A  N  F  K  K  R  E  C  V
      310                 330                 350
TCTTCTTTACCAAAGATTCCAAGGCCACGGAGAATGTGTGCAAGTGTGGCTATGCCCAGA
  F  F  T  K  D  S  K  A  T  E  N  V  C  K  C  G  Y  A  Q  S
      370                 390                 410
GCCAGCACATGGAAGGCACCCAGATCAACCAAAGTGAGAAATGGAACTACAAGAAACACA
  Q  H  M  E  G  T  Q  I  N  Q  S  E  K  W  N  Y  K  K  H  T
      430                 450                 470
CCAAGGAATTTCCTACCGACGCCTTTGGGGATATTCAGTTTGAGACACTGGGGAAGAAAG
  K  E  F  P  T  D  A  F  G  D  I  Q  F  E  T  L  G  K  K  G
      490                 510                 530
GGAAGTATATACGTCTGTCCTGCGACACGGACGCGGAAATCCTTTACGAGCTGCTGACCC
  K  Y  I  R  L  S  C  D  T  D  A  E  I  L  Y  E  L  L  T  Q
      550                 570                 590
AGCACTGGCACCTGAAAACACCCAACCTGGTCATTTCTGTGACCGGGGGCGCCAAGAACT
  H  W  H  L  K  T  P  N  L  V  I  S  V  T  G  G  A  K  N  F
      610                 630                 650
TCGCCCTGAAGCCGCGCATGCGCAAGATCTTCAGCCGGCTCATCTACATCGCGCAGTCCA
  A  L  K  P  R  M  R  K  I  F  S  R  L  I  Y  I  A  Q  S  K
      670                 690                 710
AAGGTGCTTGGATTCTCACGGGAGGCACCCATTATGGCCTGATGAAGTACATCGGGGAGG
  G  A  W  I  L  T  G  G  T  H  Y  G  L  M  K  Y  I  G  E  V
      730                 750                 770
TGGTGAGAGATAACACCATCAGCAGGAGTTCAGAGGAGAATATTGTGGCCATTGGCATAG
  V  R  D  N  T  I  S  R  S  S  E  E  N  I  V  A  I  G  I  A
      790                 810                 830
CAGCTTGGGGCATGGTCTCCAACCGGGACACCCTCATCAGGAATTGCGATGCTGAGGGCT
  A  W  G  M  V  S  N  R  D  T  L  I  R  N  C  D  A  E  G  Y
      850                 870                 890
ATTTTTTAGCCCAGTACCTTATGGATGACTTCACAAGAGATCCACTGTATATCCTGGACA
  F  L  A  Q  Y  L  M  D  D  F  T  R  D  P  L  Y  I  L  D  N
      910                 930                 950
ACAACCACACACATTTGCTGCTCGTGGACAATGGCTGTCATGGACATCCCACTGTCGAAG
  N  H  T  H  L  L  L  V  D  N  G  C  H  G  H  P  T  V  E  A
      970                 990                1010
CAAAGCTCCGGAATCAGCTAGAGAAGTATATCTCTGAGCGCACTATTCAAGATTCCAACT
  K  L  R  N  Q  L  E  K  Y  I  S  E  R  T  I  Q  D  S  N  Y
     1030                1050                1070
ATGGTGGCAAGATCCCCATTGTGTGTTTTGCCCAAGGAGGTGGAAAAGAGACTTTGAAAG
  G  G  K  I  P  I  V  C  F  A  Q  G  G  G  K  E  T  L  K  A
     1090                1110                1130
CCATCAATACCTCCATCAAAAATAAAATTCCTTGTGTGGTGGTGGAAGGCTCGGGCCAGA
  I  N  T  S  I  K  N  K  I  P  C  V  V  V  E  G  S  G  Q  I
     1150                1170                1190
TCGCTGATGTGATCGCTAGCCTGGTGGAGGTGGAGGATGCCCTGACATCTTCTGCCGTCA
  A  D  V  I  A  S  L  V  E  V  E  D  A  L  T  S  S  A  V  K
     1210                1230                1250
```

Fig. 10 / continuation 1

```
AGGAGAAGCTGGTGCGCTTTTTACCCCGCACGGTGTCCCGGCTGCCTGAGGAGGAGACTG
  E  K  L  V  R  F  L  P  R  T  V  S  R  L  P  E  E  E  T  E
     1270              1290              1310
AGAGTTGGATCAAATGGCTCAAAGAAATTCTCGAATGTTCTCACCTATTAACAGTTATTA
  S  W  I  K  W  L  K  E  I  L  E  C  S  H  L  L  T  V  I  K
     1330              1350              1370
AAATGGAAGAAGCTGGGGATGAAATTGTGAGCAATGCCATCTCCTACGCTCTATACAAAG
  M  E  E  A  G  D  E  I  V  S  N  A  I  S  Y  A  L  Y  K  A
     1390              1410              1430
CCTTCAGCACCAGTGAGCAAGACAAGGATAACTGGAATGGGCAGCTGAAGCTTCTGCTGG
  F  S  T  S  E  Q  D  K  D  N  W  N  G  Q  L  K  L  L  L  E
     1450              1470              1490
AGTGGAACCAGCTGGACTTAGCCAATGATGAGATTTTCACCAATGACCGCCGATGGGAGA
  W  N  Q  L  D  L  A  N  D  E  I  F  T  N  D  R  R  W  E  K
     1510              1530              1550
AGAGCAAACCGAGGCTCAGAGACACAATAATCCAGGTCACATGGCTGGAAAATGGTAGAA
  S  K  P  R  L  R  D  T  I  I  Q  V  T  W  L  E  N  G  R  I
     1570              1590              1610
TCAAGGTTGAGAGCAAAGATGTGACTGACGGCAAAGCCTCTTCTCATATGCTGGTGGTTC
  K  V  E  S  K  D  V  T  D  G  K  A  S  S  H  M  L  V  V  L
     1630              1650              1670
TCAAGTCTGCTGACCTTCAAGAAGTCATGTTTACGGCTCTCATAAAGGACAGACCCAAGT
  K  S  A  D  L  Q  E  V  M  F  T  A  L  I  K  D  R  P  K  F
     1690              1710              1730
TTGTCCGCCTCTTTCTGGAGAATGGCTTGAACCTACGGAAGTTTCTCACCCATGATGTCC
  V  R  L  F  L  E  N  G  L  N  L  R  K  F  L  T  H  D  V  L
     1750              1770              1790
TCACTGAACTCTTCTCCAACCACTTCAGCACGCTTGTGTACCGGAATCTGCAGATCGCCA
  T  E  L  F  S  N  H  F  S  T  L  V  Y  R  N  L  Q  I  A  K
     1810              1830              1850
AGAATTCCTATAATGATGCCCTCCTCACGTTTGTCTGGAAACTGGTTGCGAACTTCCGAA
  N  S  Y  N  D  A  L  L  T  F  V  W  K  L  V  A  N  F  R  R
     1870              1890              1910
GAGGCTTCCGGAAGGAAGACAGAAATGGCCGGGACGAGATGGACATAGAACTCCACGACG
  G  F  R  K  E  D  R  N  G  R  D  E  M  D  I  E  L  H  D  V
     1930              1950              1970
TGTCTCCTATTACTCGGCACCCCCTGCAAGCTCTCTTCATCTGGGCCATTCTTCAGAATA
  S  P  I  T  R  H  P  L  Q  A  L  F  I  W  A  I  L  Q  N  K
     1990              2010              2030
AGAAGGAACTCTCCAAAGTCATTTGGGAGCAGACCAGGGGCTGCACTCTGGCAGCCCTGG
  K  E  L  S  K  V  I  W  E  Q  T  R  G  C  T  L  A  A  L  G
     2050              2070              2090
GAGCCAGCAAGCTTCTGAAGACTCTGGCCAAAGTGAAGAACGACATCAATGCTGCTGGGG
  A  S  K  L  L  K  T  L  A  K  V  K  N  D  I  N  A  A  G  E
     2110              2130              2150
AGTCCGAGGAGCTGGCTAATGAGTACGAGACCCGGGCTGTTGGTGAGTCCACAGTGTGGA
  S  E  E  L  A  N  E  Y  E  T  R  A  V  G  E  S  T  V  W  N
     2170              2190              2210
ATGCTGTGGTGGGCGCGGATCTGCCATGTGGCACAGACATTGCCAGCGGCACTCATAGAC
  A  V  V  G  A  D  L  P  C  G  T  D  I  A  S  G  T  H  R  P
     2230              2250              2270
CAGATGGTGGAGAGCTGTTCACTGAGTGTTACAGCAGCGATGAAGACTTGGCAGAACAGC
  D  G  G  E  L  F  T  E  C  Y  S  S  D  E  D  L  A  E  Q  L
     2290              2310              2330
TGCTGGTCTATTCCTGTGAAGCTTGGGGTGGAAGCAACTGTCTGGAGCTGGCGGTGGAGG
  L  V  Y  S  C  E  A  W  G  G  S  N  C  L  E  L  A  V  E  A
     2350              2370              2390
CCACAGACCAGCATTTCATCGCCCAGCCTGGGGTCCAGAATTTTCTTTCTAAGCAATGGT
  T  D  Q  H  F  I  A  Q  P  G  V  Q  N  F  L  S  K  Q  W  Y
     2410              2430              2450
ATGGAGAGATTTCCCGAGACACCAAGAACTGGAAGATTATCCTGTGTCTGTTTATTATAC
  G  E  I  S  R  D  T  K  N  W  K  I  I  L  C  L  F  I  I  P
```

Fig. 10 / continuation 2

```
     2470                2490                2510
CCTTGGTGGGCTGTGGCTTTGTATCATTTAGGAAGAAACCTGTCGACAAGCACAAGAAGC
  L  V  G  C  G  F  V  S  F  R  K  K  P  V  D  K  H  K  K  L
     2530                2550                2570
TGCTTTGGTACTATGTGGCGTTCTTCACCTCCCCCTTCGTGGTCTTCTCCTGGAATGTGG
  L  W  Y  Y  V  A  F  F  T  S  P  F  V  V  F  S  W  N  V  V
     2590                2610                2630
TCTTCTACATCGCCTTCCTCCTGCTGTTTGCCTACGTGCTGCTCATGGATTTCCATTCGG
  F  Y  I  A  F  L  L  L  F  A  Y  V  L  L  M  D  F  H  S  V
     2650                2670                2690
TGCCACACCCCCCCGAGCTGGTCCTGTACTCGCTGGTCTTTGTCCTCTTCTGTGATGAAG
  P  H  P  P  E  L  V  L  Y  S  L  V  F  V  L  F  C  D  E  V
     2710                2730                2750
TGAGACAGGGCCGGCCGGCTGCTCCCAGTGCGGGGCCCGCCAAGCCCACGCCCACCCGGA
  R  Q  G  R  P  A  A  P  S  A  G  P  A  K  P  T  P  T  R  N
     2770                2790                2810
ACTCCATCTGGCCCGCAAGCTCCACACGCAGCCCCGGTTCCCGCTCACGCCACTCCTTCC
  S  I  W  P  A  S  S  T  R  S  P  G  S  R  S  R  H  S  F  H
     2830                2850                2870
ACACTTCCCTGCAAGCTGAGGGTGCCAGCTCTGGCCTTGGCCAGCCCAGAAAGGGGTGGA
  T  S  L  Q  A  E  G  A  S  S  G  L  G  Q  P  R  K  G  W  T
     2890                2910                2930
CATTTAAAAATCTGGAAATGGTTGATATTTCCAAGCTGCTGATGTCCCTCTCTGTCCCTT
  F  K  N  L  E  M  V  D  I  S  K  L  L  M  S  L  S  V  P  F
     2950                2970                2990
TCTGTACGCAGTGGTACGTAAATGGGGTGAATTATTTTACTGACCTGTGGAATGTGATGG
  C  T  Q  W  Y  V  N  G  V  N  Y  F  T  D  L  W  N  V  M  D
     3010                3030                3050
ACACGCTGGGGCTTTTTTACTTCATAGCAGGAATTGTATTTCGGCAAGGGATCCTTAGGC
  T  L  G  L  F  Y  F  I  A  G  I  V  F  R  Q  G  I  L  R  Q
     3070                3090                3110
AGAATGAGCAGCGCTGGAGGTGGATATTCCGTTCGGTCATCTACGAGCCCTACCTGGCCA
  N  E  Q  R  W  R  W  I  F  R  S  V  I  Y  E  P  Y  L  A  M
     3130                3150                3170
TGTTCGGCCAGGTGCCCAGTGACGTGGATGGTACCACGTATGACTTTGCCCACTGCACCT
  F  G  Q  V  P  S  D  V  D  G  T  T  Y  D  F  A  H  C  T  F
     3190                3210                3230
TCACTGGGAATGAGTCCAAGCCACTGTGTGTGGAGCTGGATGAGCACAACCTGCCCCGGT
  T  G  N  E  S  K  P  L  C  V  E  L  D  E  H  N  L  P  R  F
     3250                3270                3290
TCCCCGAGTGGATCACCATCCCCCTGGTGTGCATCTACATGTTATCCACCAACATCCTGC
  P  E  W  I  T  I  P  L  V  C  I  Y  M  L  S  T  N  I  L  L
     3310                3330                3350
TGGTCAACCTGCTGGTCGCCATGTTTGGCTACACGGTGGGCACCGTCCAGGAGAACAATG
  V  N  L  L  V  A  M  F  G  Y  T  V  G  T  V  Q  E  N  N  D
     3370                3390                3410
ACCAGGTCTGGAAGTTCCAGAGGTACTTCCTGGTGCAGGAGTACTGCAGCCGCCTCAATA
  Q  V  W  K  F  Q  R  Y  F  L  V  Q  E  Y  C  S  R  L  N  I
     3430                3450                3470
TCCCCTTCCCCTTCATCGTCTTCGCTTACTTCTACATGGTGGTGAAGAAGTGCTTCAAGT
  P  F  P  F  I  V  F  A  Y  F  Y  M  V  V  K  K  C  F  K  C
     3490                3510                3530
GTTGCTGCAAGGAGAAAAACATGGAGTCTTCTGTCTGCTGTGAGTGGTTTATCCATGTGT
  C  C  K  E  K  N  M  E  S  S  V  C  C  E  W  F  I  H  V  Y
     3550                3570                3590
ACTTGGGATCAGAAGCAGCGATTAATTTCAGGGAAGGATGCCTGCATCCAGTGATTGGAA
  L  G  S  E  A  A  I  N  F  R  E  G  C  L  H  P  V  I  G  S
     3610                3630                3650
GCTGGACCCCAGGCTGGCTGGTCTGGACATCCACACGCATTCTCACATGCAGTGCCGGCT
  W  T  P  G  W  L  V  W  T  S  T  R  I  L  T  C  S  A  G  W
     3670                3690                3710
GGCCAGCAGCAGGGAGTCTCAGTGTCACCACACATAGCAGCTGGGTTCCTGCAAAAAGCA
```

Fig. 10 / continuation 3

```
  P  A  A  G  S  L  S  V  T  T  H  S  S  W  V  P  A  K  S  S
     3730              3750              3770
GCAAGTCACAGGCCCACCCAGACAGAACGGGTAGAGAATGTGACTCTGCTTCTGGGTGGG
  K  S  Q  A  H  P  D  R  T  G  R  E  C  D  S  A  S  G  W  E
     3790              3810              3830
AAGGACAGCCTGCCCGGTGGGTGGAAGAATCCGTGGCCCTGTTTGGCCATCGTGGCCCTG
  G  Q  P  A  R  W  V  E  E  S  V  A  L  F  G  H  R  G  P  V
     3850              3870              3890
TTTGGCCACCTACCACTCTAGGCATCACTGAGCTGAATGCGCCGGTCCTCTGA
  W  P  P  T  T  L  G  I  T  E  L  N  A  P  V  L  *
```

MVGGCRWTEDVEPAEVKEKMSFRAARLSMRNRRNDTLDSTRTLYSSASRSTDLSYSESASFYAAFRTQTCPIMASWDLVNFIQANF
KKRECVFFTKDSKATENVCKCGYAQSQHMEGTQINQSEKWNYKKHTKEFPTDAFGDIQFETLGKKGKYIRLSCDTDAEILYELLTQ
HWHLKTPNLVISVTGGAKNFALKPRMRKIFSRLIYIAQSKGAWILTGGTHYGLMKYIGEVVRDNTISRSSEENIVAIGIAAWGMVS
NRDTLIRNCDAEGYFLAQYLMDDFTRDPLYILDNNHTHLLLVDNGCHGHPTVEAKLRNQLEKYISERTIQDSNYGGKIPIVCFAQG
GGKETLKAINTSIKNKIPCVVVEGSGQIADVIASLVEVEDALTSSAVKEKLVRFLPRTVSRLPEEETESWIKWLKEILECSHLLTV
IKMEEAGDEIVSNAISYALYKAFSTSEQDKDNWNGQLKLLLEWNQLDLANDEIFTNDRRWEKSKPRLRDTIIQVTWLENGRIKVES
KDVTDGKASSHMLVVLKSADLQEVMFTALIKDRPKFVRLFLENGLNLRKFLTHDVLTELFSNHFSTLVYRNLQIAKNSYNDALLTF
VWKLVANFRRGFRKEDRNGRDEMDIELHDVSPITRHPLQALFIWAILQNKKELSKVIWEQTRGCTLAALGASKLLKTLAKVKNDIN
AAGESEELANEYETRAVGESTVWNAVVGADLPCGTDIASGTHRPDGGELFTECYSSDEDLAEQLLVYSCEAWGGSNCLELAVEATD
QHFIAQPGVQNFLSKQWYGEISRDTKNWKIILCLFIIPLVGCGFVSFRKKPVDKHKKLLMYYVAFFTSPFVVFSWNVVFYIAFLLL
FAYVLLMDFHSVPHPPELVLYSLVFVLFCDEVRQGRPAAPSAGPAKPTPTRNSIWPASSTRSPGSRSRHSFHTSLQAEGASSGLGQ
PRKGWTFKNLEMVDISKLLMSLSVPFCTQWYVNGVNYFTDLWNVMDTLGLFYFIAGIVFRQGILRQNEQRWRWIFRSVIYEPYLAM
FGQVPSDVDGTTYDFAHCTFTGNESKPLCVELDEHNLPRFPEWITIPLVCIYMLSTNILLVNLLVAMFGYTVGTVQENNDQVWKFQ
RYFLVQEYCSRLNIPFPFIVFAYFYMVVKKCFKCCCKEKNMESSVCCEWFIHVYLGSEAAINFREGCLHPVIGSWTPGWLVWTSTR
ILTCSAGWPAAGSLSVTTHSSWVPAKSSKSQAHPDRTGRECDSASGWEGQPARWVEESVALFGHRGPVWPPTTLGITELNAPVL

B.

```
                                                      Q  L
     2290              2310              2330
TGCTGGTCTATTCCTGTGAAGCTTGGGGTGGAAGCAACTGTCTGGAGCTGGCGGTGGAGG
  L  V  Y  S  C  E  A  W  G  G  S  N  C  L  E  L  A  V  E  A
     2350              2370              2390
CCACAGACCAGCATTTCATCGCCCAGCCTGGGGTCCAGAATTTTCTTTCTAAGCAATGGT
  T  D  Q  H  F  I  A  Q  P  G  V  Q  N  F  L  S  K  Q  W  Y
     2410              2430              2450
ATGGAGAGATTTCCCGAGACACCAAGAACTGGAAGATTATCCTGTGTCTGTTTATTATAC
  G  E  I  S  R  D  T  K  N  W  K  I  I  L  C  L  F  I  I  P
     2470              2490              2510
CCTTGGTGGGCTGTGGCTTTGTATCATTTAGGAAGAAACCTGTCGACAAGCACAAGAAGC
  L  V  G  C  G  F  V  S  F  R  K  K  P  V  D  K
```

Figure 11:

a.) Trp10b cDNA and derived amino acid sequence

```
        10                  30                  50
ATGAAATCCTTCCTTCCTGTCCACACCATCGTGCTTATCAGGGAGAATGTGTGCAAGTGT
M  K  S  F  L  P  V  H  T  I  V  L  I  R  E  N  V  C  K  C
        70                  90                 110
GGCTATGCCCAGAGCCAGCACATGGAAGGCACCCAGATCAACCAAAGTGAGAAATGGAAC
G  Y  A  Q  S  Q  H  M  E  G  T  Q  I  N  Q  S  E  K  W  N
       130                 150                 170
TACAAGAAACACACCAAGGAATTTCCTACCGACGCCTTTGGGGATATTCAGTTTGAGACA
Y  K  K  H  T  K  E  F  P  T  D  A  F  G  D  I  Q  F  E  T
       190                 210                 230
CTGGGGAAGAAAGGGAAGTATATACGTCTGTCCTGCGACACGGACGCGGAAATCCTTTAC
L  G  K  K  G  K  Y  I  R  L  S  C  D  T  D  A  E  I  L  Y
       250                 270                 290
GAGCTGCTGACCCAGCACTGGCACCTGAAAACACCCAACCTGGTCATTTCTGTGACCGGG
E  L  L  T  Q  H  W  H  L  K  T  P  N  L  V  I  S  V  T  G
       310                 330                 350
GGCGCCAAGAACTTCGCCCTGAAGCCGCGCATGCGCAAGATCTTCAGCCGGCTCATCTAC
G  A  K  N  F  A  L  K  P  R  M  R  K  I  F  S  R  L  I  Y
       370                 390                 410
ATCGCGCAGTCCAAAGGTGCTTGGATTCTCACGGGAGGCACCCATTATGGCCTGATGAAG
I  A  Q  S  K  G  A  W  I  L  T  G  G  T  H  Y  G  L  M  K
       430                 450                 470
TACATCGGGGAGGTGGTGAGAGATAACACCATCAGCAGGAGTTCAGAGGAGAATATTGTG
Y  I  G  E  V  V  R  D  N  T  I  S  R  S  S  E  E  N  I  V
       490                 510                 530
GCCATTGGCATAGCAGCTTGGGGCATGGTCTCCAACCGGGACACCCTCATCAGGAATTGC
A  I  G  I  A  A  W  G  M  V  S  N  R  D  T  L  I  R  N  C
       550                 570                 590
GATGCTGAGGGCTATTTTTTAGCCCAGTACCTTATGGATGACTTCACAAGAGATCCACTG
D  A  E  G  Y  F  L  A  Q  Y  L  M  D  D  F  T  R  D  P  L
       610                 630                 650
TATATCCTGGACAACAACCACACACATTTGCTGCTCGTGGACAATGGCTGTCATGGACAT
Y  I  L  D  N  N  H  T  H  L  L  L  V  D  N  G  C  H  G  H
       670                 690                 710
CCCACTGTCGAAGCAAAGCTCCGGAATCAGCTAGAGAAGTATATCTCTGAGCGCACTATT
P  T  V  E  A  K  L  R  N  Q  L  E  K  Y  I  S  E  R  T  I
       730                 750                 770
CAAGATTCCAACTATGGTGGCAAGATCCCCATTGTGTGTTTTGCCCAAGGAGGTGGAAAA
Q  D  S  N  Y  G  G  K  I  P  I  V  C  F  A  Q  G  G  G  K
       790                 810                 830
GAGACTTTGAAAGCCATCAATACCTCCATCAAAAATAAAATTCCTTGTGTGGTGGTGGAA
E  T  L  K  A  I  N  T  S  I  K  N  K  I  P  C  V  V  V  E
       850                 870                 890
GGCTCGGGCCAGATCGCTGATGTGATCGCTAGCCTGGTGGAGGTGGAGGATGCCCTGACA
G  S  G  Q  I  A  D  V  I  A  S  L  V  E  V  E  D  A  L  T
       910                 930                 950
TCTTCTGCCGTCAAGGAGAAGCTGGTGCGCTTTTTACCCCGCACGGTGTCCCGGCTGCCT
S  S  A  V  K  E  K  L  V  R  F  L  P  R  T  V  S  R  L  P
       970                 990                1010
GAGGAGGAGACTGAGAGTTGGATCAAATGGCTCAAAGAAATTCTCGAATGTTCTCACCTA
E  E  E  T  E  S  W  I  K  W  L  K  E  I  L  E  C  S  H  L
      1030                1050                1070
TTAACAGTTATTAAAATGGAAGAAGCTGGGGATGAAATTGTGAGCAATGCCATCTCCTAC
L  T  V  I  K  M  E  E  A  G  D  E  I  V  S  N  A  I  S  Y
      1090                1110                1130
GCTCTATACAAAGCCTTCAGCACCAGTGAGCAAGACAAGGATAACTGGAATGGGCAGCTG
A  L  Y  K  A  F  S  T  S  E  Q  D  K  D  N  W  N  G  Q  L
```

Fig. 11 (Continuation)

```
      2410                2430                2450
AGAAACTTAGGACCCAAGATTATAATGCTGCAGAGGATGCTGATCGATGTGTTCTTCTTC
R  N  L  G  P  K  I  I  M  L  Q  R  M  L  I  D  V  F  F  F
      2470                2490                2510
CTGTTCCTCTTTGCGGTGTGGATGGTGGCCTTTGGCGTGGCCAGGCAAGGGATCCTTAGG
L  F  L  F  A  V  W  M  V  A  F  G  V  A  R  Q  G  I  L  R
      2530                2550                2570
CAGAATGAGCAGCGCTGGAGGTGGATATTCCGTTCGGTCATCTACGAGCCCTACCTGGCC
Q  N  E  Q  R  W  R  W  I  F  R  S  V  I  Y  E  P  Y  L  A
      2590                2610                2630
ATGTTCGGCCAGGTGCCCAGTGACGTGGATGGTACCACGTATGACTTTGCCCACTGCACC
M  F  G  Q  V  P  S  D  V  D  G  T  T  Y  D  F  A  H  C  T
      2650                2670                2690
TTCACTGGGAATGAGTCCAAGCCACTGTGTGTGGAGCTGGATGAGCACAACCTGCCCCGG
F  T  G  N  E  S  K  P  L  C  V  E  L  D  E  H  N  L  P  R
      2710                2730                2750
TTCCCCGAGTGGATCACCATCCCCCTGGTGTGCATCTACATGTTATCCACCAACATCCTG
F  P  E  W  I  T  I  P  L  V  C  I  Y  M  L  S  T  N  I  L
      2770                2790                2810
CTGGTCAACCTGCTGGTCGCCATGTTTGGCTACACGGTGGGCACCGTCCAGGAGAACAAT
L  V  N  L  L  V  A  M  F  G  Y  T  V  G  T  V  Q  E  N  N
      2830                2850                2870
GACCAGGTCTGGAAGTTCCAGAGGTACTTCCTGGTGCAGGAGTACTGCAGCCGCCTCAAT
D  Q  V  W  K  F  Q  R  Y  F  L  V  Q  E  Y  C  S  R  L  N
      2890                2910                2930
ATCCCCTTCCCCTTCATCGTCTTCGCTTACTTCTACATGGTGGTGAAGAAGTGCTTCAAG
I  P  F  P  F  I  V  F  A  Y  F  Y  M  V  V  K  K  C  F  K
      2950                2970                2990
TGTTGCTGCAAGGAGAAAAACATGGAGTCTTCTGTCTGCTGTTTCAAAAATGAAGACAAT
C  C  C  K  E  K  N  M  E  S  S  V  C  C  F  K  N  E  D  N
      3010                3030                3050
GAGACTCTGGCATGGGAGGGTGTCATGAAGGAAAACTACCTTGTCAAGATCAACACAAAA
E  T  L  A  W  E  G  V  M  K  E  N  Y  L  V  K  I  N  T  K
      3070                3090                3110
GCCAACGACACCTCAGAGGAAATGAGGCATCGATTTAGACAACTGGATACAAAGCTTAAT
A  N  D  T  S  E  E  M  R  H  R  F  R  Q  L  D  T  K  L  N
      3130                3150
GATCTCAAGGGTCTACTGAAAGAGATTGCTAATAAAATCAAATAG
D  L  K  G  L  L  K  E  I  A  N  K  I  K  *
```

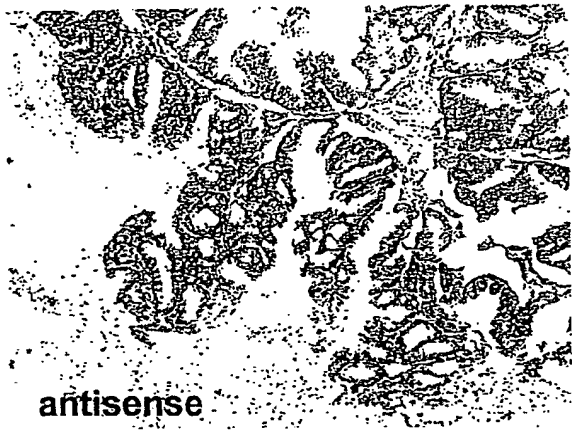
b.) Trp10 protein:

```
MKSFLPVHTIVLIRENVCKCGYAQSQHMEGTQINQSEKWNYKKHTKEFPTDAFGDIQFETLGKKGKYIRLSCDTDAEILY
ELLTQHWHLKTPNLVISVTGGAKNFALKPRMRKIFSRLIYIAQSKGAWILTGGTHYGLMKYIGEVVRDNTISRSSEENIV
AIGIAAWGMVSNRDTLIRNCDAEGYFLAQYLMDDFTRDPLYILDNNHTHLLLVDNGCHGHPTVEAKLRNQLEKYISERTI
QDSNYGGKIPIVCFAQGGGKETLKAINTSIKNKIPCVVVEGSGQIADVIASLVEVEDALTSSAVKEKLVRFLPRTVSRLP
EEETESWIKWLKEILECSHLLTVIKMEEAGDEIVSNAISYALYKAFSTSKQDKDNWNGQLKLLLEWNQLDLANDEIFTND
RRWESADLQEVMFTALIKDRPKFVRLFLENGLNLRKFLTHDVLTELFSNHFSTLVYRNLQIAKNSYNDALLTFVWKLVAN
FRRGFRKEDRNGRDEMDIELHDVSPITRHPLQALFIWAILQNKKELSKVIWEQTRGCTLAALGASKLLKTLAKVKNDINA
AGESEELANEYETRAVELFTECYSSDEDLAEQLLVYSCEAWGGSNCLELAVEATDQHFIAQPGVQNFLSKQWYGEISRDT
KNWKIILCLFIIPLVGCGFVSFRKKPVDKHKKLLWYYVAFFTSPFVVFSWNVVFYIAFLLLFAYVLLMDFHSVPHPPELV
LYSLVFVLFCDEVRQWYVNGVNYFTDLWNVMDTLGLFYFIAGIVFRLHSSNKSSLYSGRVIFCLDYIIFTLRLIHIFTVS
RNLGPKIIMLQRMLIDVFFFLFLFAVWMVAFGVARQGILRQNEQRWRWIFRSVIYEPYLAMFGQVPSDVDGTTYDFAHCT
FTGNESKPLCVELDEHNLPRFPEWITIPLVCIYMLSTNILLVNLLVAMFGYTVGTVQENNDQVWKFQRYFLVQEYCSRLN
IPFPFIVFAYFYMVVKKCFKCCCKEKNMESSVCCFKNEDNETLAWEGVMKENYLVKINTKANDTSEEMRHRFRQLDTKLN
DLKGLLKEIANKIK
```

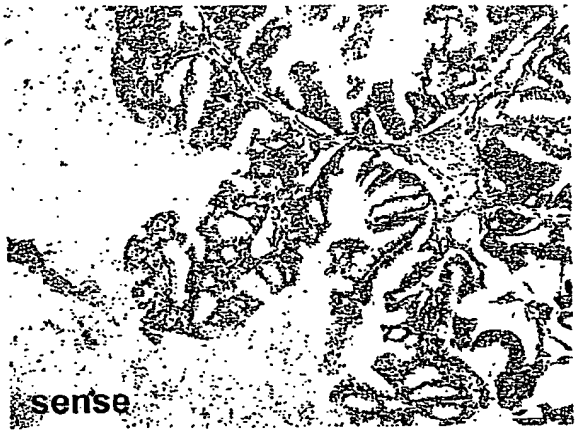
The Trp8 Gene is expre d in endometrial or uteri
in normal endometrium

Endometrial cancer:

A



B



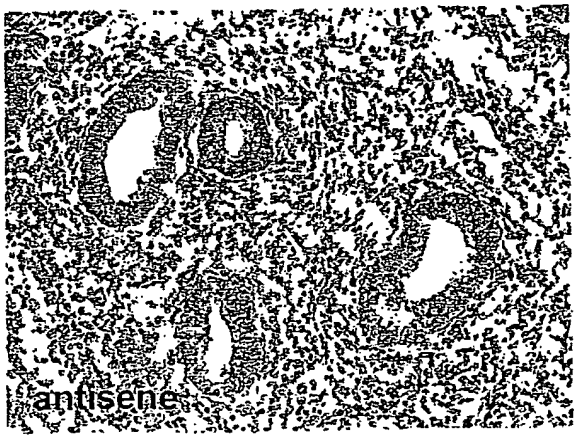
C

D

Endometrium:

E

F



Expression of human Trp 9 and Trp 10

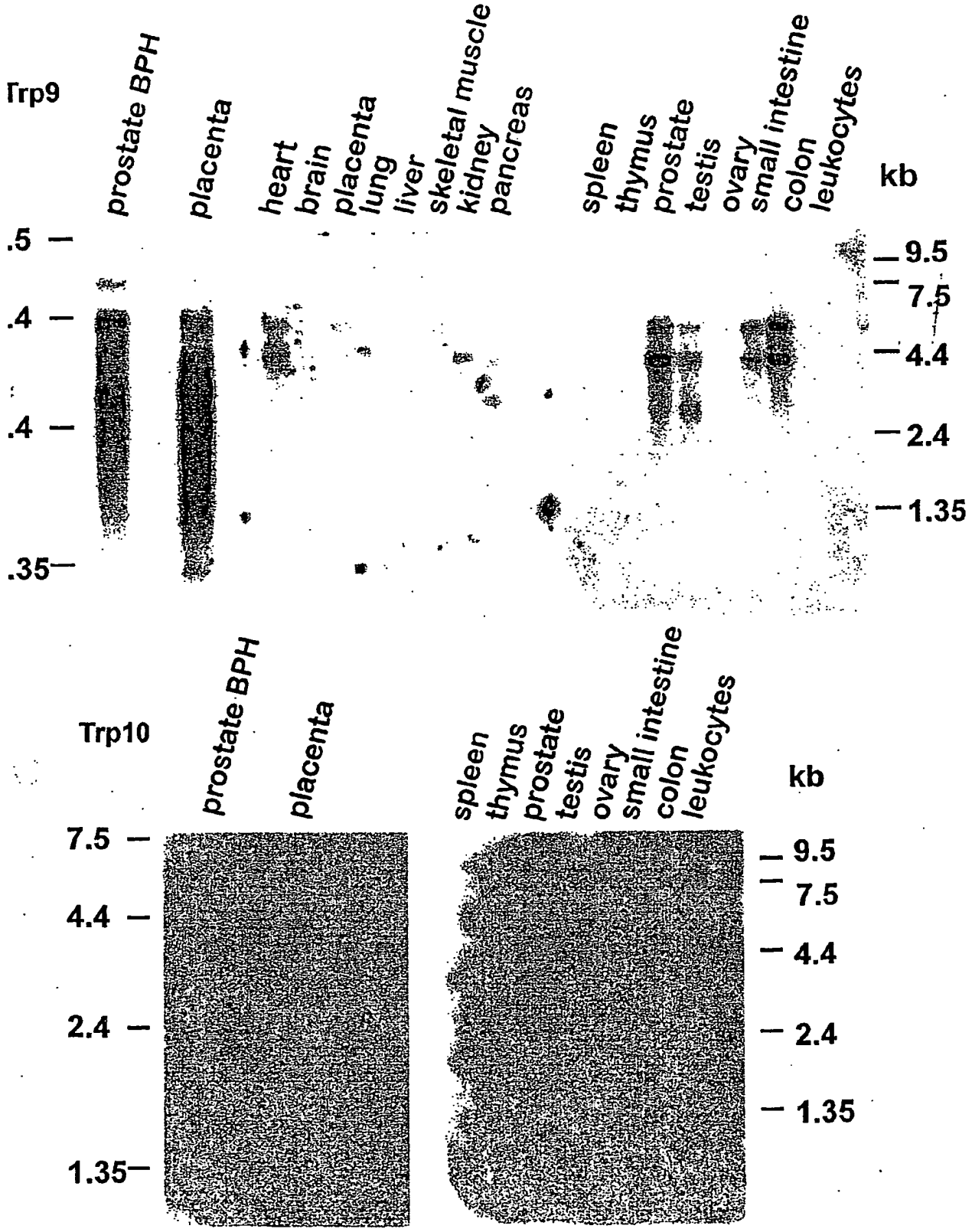
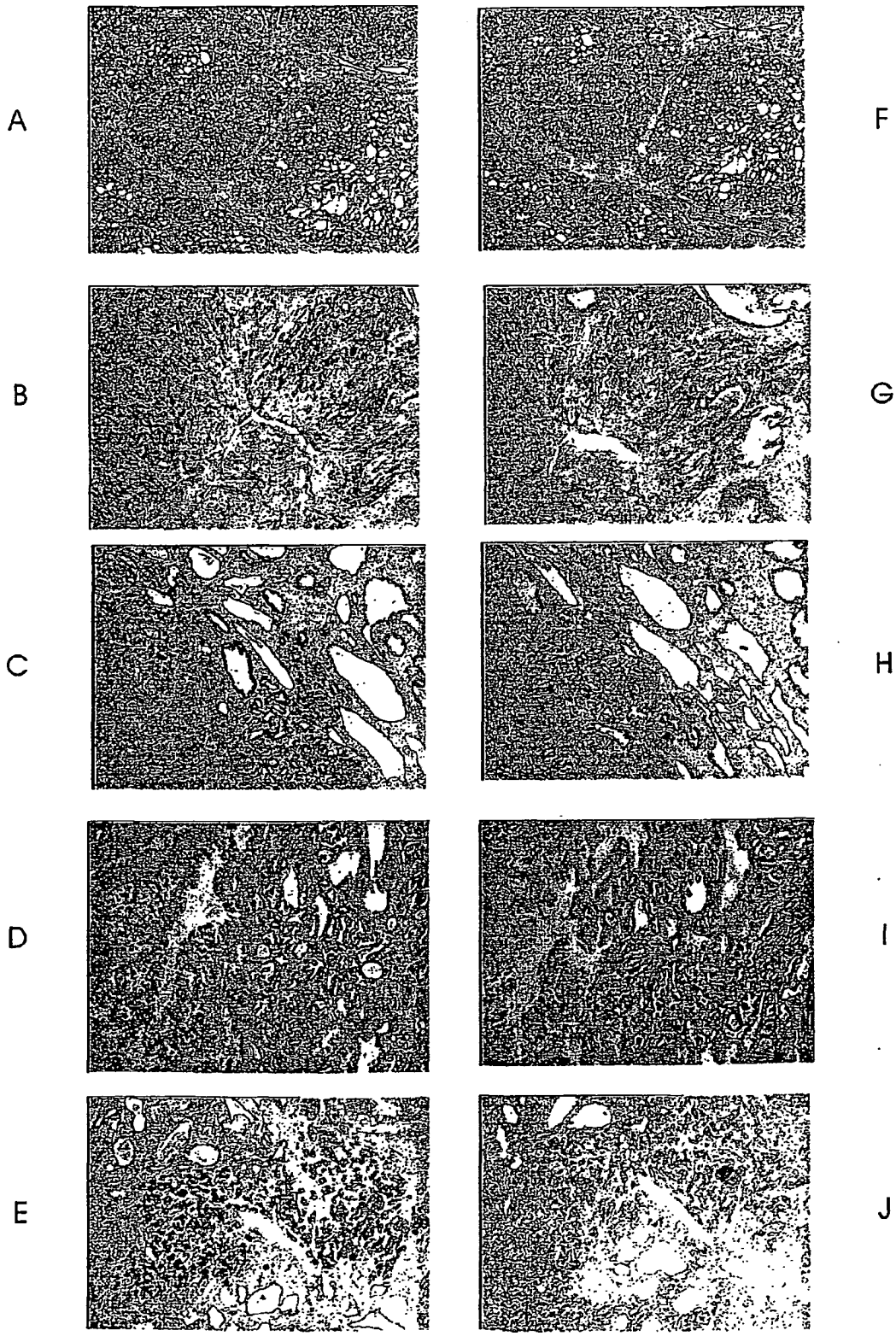
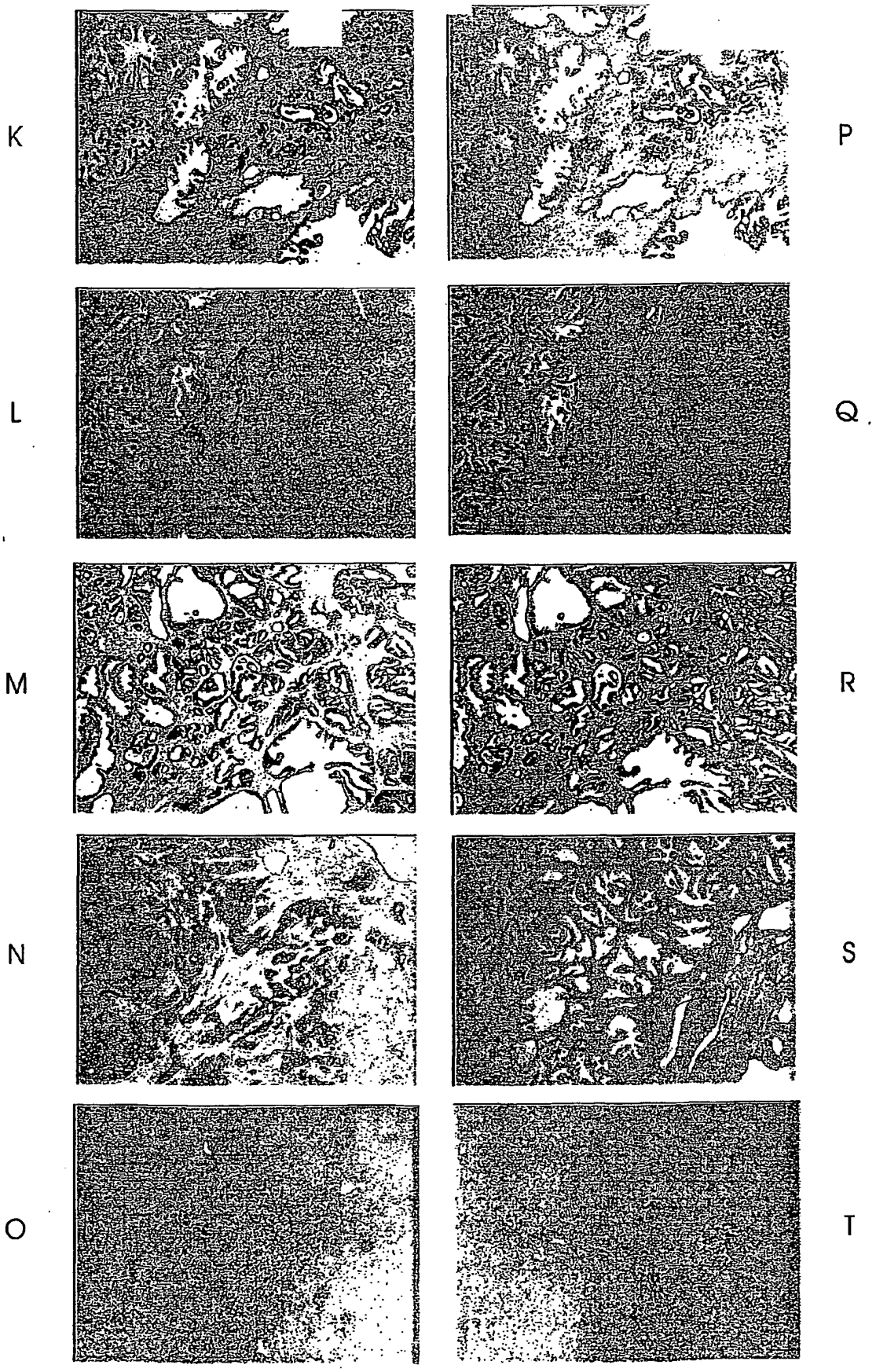


Fig. 14

Expression of Trp10 transcripts and Trp10-antisense transcripts in human prostate cancer and in malignant melanoma




K
P
L
Q
M
R
N
S
O
T

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 February 2002 (07.02.2002)

PCT

(10) International Publication Number
WO 02/010382 A3

(51) International Patent Classification[7]: C12N 15/12, 15/11, 9/00, C07K 14/47, C12Q 1/68, G01N 33/577, A61K 31/713, C07K 14/705

(21) International Application Number: PCT/EP01/08309

(22) International Filing Date: 18 July 2001 (18.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/221,513 28 July 2000 (28.07.2000) US

(71) Applicant and
(72) Inventor: **WISSENBACH, Ulrich** [DE/DE]; Institut für Pharmakologie und Toxikologie der Uni, versität des Saarlandes, 66421 Homburg (DE).

(74) Agent: **HUBER, Bernard**; Huber & Schüssler, Truderinger Str. 246, 81825 München (DE).

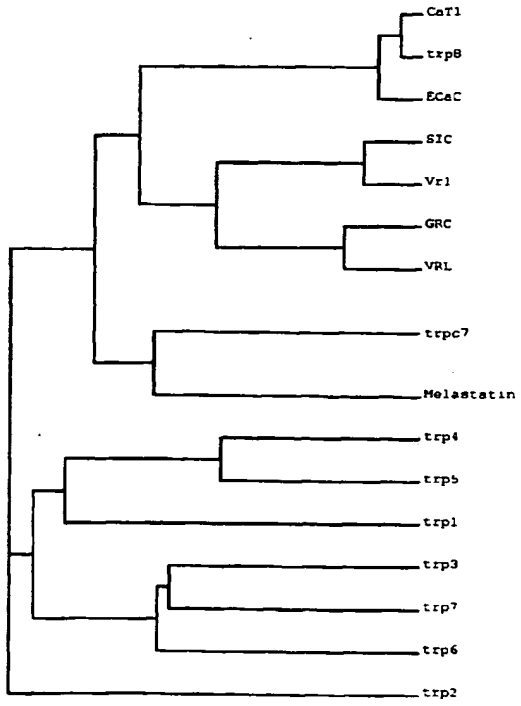
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

[Continued on next page]

(54) Title: TRP8, TRP9 AND TRP10, MARKERS FOR CANCER



(57) Abstract: The present invention relates to gene expression in normal cells and cells of malignant tumors and particularly to novel markers associated with cancer, Trp8, Trp9 and Trp10, and the genes encoding Trp8, Trp9 and Trp10. Also provided are vectors, host cells, antibodies, and recombinant methods for producing these human proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating a tumor.

WO 02/010382 A3

(88) **Date of publication of the international search report:**
9 October 2003

(15) **Information about Correction:**
Previous Correction:
see PCT Gazette No. 38/2002 of 19 September 2002, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

rnational Application No
PCT/EP 01/08309

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C12N15/11 C12N9/00 C07K14/47 C12Q1/68
G01N33/577 A61K31/713 C07K14/705

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)
BIOSIS, EPO-Internal, SEQUENCE SEARCH, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 09166 A (SHAPERO MICHAEL H ;DENDREON CORP (US); LAUS REINER (US); TSAVALER) 25 February 1999 (1999-02-25) see SEQID14 + 15, pages 2,3, 28,29, Example 4 table 3 ---	1-10, 12-17, 23,29-31
X	WO 00 40614 A (BETH ISRAEL HOSPITAL ;SCHARENBERG ANDREW M (US)) 13 July 2000 (2000-07-13) see seqid31 + 32, page 11, first paragraph, page 44, lines 13-15 ---	1-10,12, 31

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the International filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the International filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the International search report
6 March 2003	13. 03. 2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Holtorf, S

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MULLER D ET AL: "Molecular cloning, tissue distribution, and chromosomal mapping of the human epithelial Ca2+ channel (ECAC1)." GENOMICS, vol. 67, no. 1, 1 July 2000 (2000-07-01), pages 48-53, XP002222953 ISSN: 0888-7543 the whole document	1
X	WO 98 15657 A (ABBOTT LAB) 16 April 1998 (1998-04-16) the whole document	1-12, 29-31
X	WO 98 37093 A (CORIXA CORP) 27 August 1998 (1998-08-27) the whole document	1-12, 29-31
A	TSAVALER LARISA ET AL: "TRP-P8, a novel prostate-specific gene, is upregulated in prostate cancer and other malignancies and shares high homology with TRP calcium channel proteins." PROCEEDINGS OF THE AMERICAN ASSOCIATION FOR CANCER RESEARCH ANNUAL, no. 41, March 2000 (2000-03), page 694 XP008011242 91st Annual Meeting of the American Association for Cancer Research.;San Francisco, California, USA; April 01-05, 2000, March, 2000 ISSN: 0197-016X the whole document	
A	HARTENECK C ET AL: "FROM WORM TO MAN: THREE SUBFAMILIES OF TRP CHANNELS" TRENDS IN NEUROSCIENCE, ELSEVIER, AMSTERDAM, NL, vol. 23, no. 4, April 2000 (2000-04), pages 159-166, XP001012870 ISSN: 0166-2236	
P,X	WO 01 14423 A (SMITHKLINE BEECHAM PLC) 1 March 2001 (2001-03-01) see SEQid1 + 2; see example 1	1-9,31

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WISSENBACH ULRICH ET AL: "Expression of CaT-like, a novel calcium-selective channel, correlates with the malignancy of prostate cancer." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 276, no. 22, 1 June 2001 (2001-06-01), pages 19461-19468, XP002222954 ISSN: 0021-9258 the whole document	1-9,13, 14, 16-19, 21-23,29
P,X	WO 01 04303 A (HEDIGER MATTHIAS A) 18 January 2001 (2001-01-18) see SEQID1 + 2 the whole document	1-5
P,X	WO 01 42467 A (MILLENNIUM PREDICTIVE MEDICINE) 14 June 2001 (2001-06-14) see SEQID 4615	1
E	WO 01 51633 A (FANGER GARY RICHARD ;HARLOCKER SUSAN L (US); MEAGHER MADELEINE JOY) 19 July 2001 (2001-07-19) see SEQID764, example 3, claims	1
E	WO 02 14361 A (AGENSYS INC) 21 February 2002 (2002-02-21) see SEQID1479, examples 1-4 the whole document	1-10, 13-23
E	WO 02 00722 A (SILOS SANTIAGO INMACULADA ;CURTIS RORY A J (US); MILLENNIUM PHARM) 3 January 2002 (2002-01-03) see SEQID4	1-5
E	WO 01 68857 A (CURTIS RORY A J ;COOK WILLIAM JAMES (US); MILLENNIUM PHARM INC (US) 20 September 2001 (2001-09-20) see SEQID1, examples	1-5
E	WO 01 53348 A (SQUIBB BRISTOL MYERS CO ;GAUGHAN GLEN T (US); RAMANATHAN CHANDRA S) 26 July 2001 (2001-07-26) see SEQID5 the whole document	1
E	WO 01 62794 A (LORA JOSE M ;CURTIS RORY A J (US); GLUCKSMANN MARIA ALEXANDRA (US)) 30 August 2001 (2001-08-30) the whole document	1-9
E	WO 02 30268 A (EOS BIOTECHNOLOGY INC) 18 April 2002 (2002-04-18) see SEQID53	1,6

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	BOEDDING MATTHIAS ET AL: "The recombinant human TRPV6 channel functions as Ca2+ sensor in human embryonic kidney and rat basophilic leukemia cells." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 277, no. 39, 27 September 2002 (2002-09-27), pages 36656-36664, XP002222955 September 27, 2002 ISSN: 0021-9258 the whole document -----	

INTERNATIONAL SEARCH REPORT

nternational application No.
PCT/EP 01/08309

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 24-28 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.: 12 partially
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [X] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[X] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-12, 29-31 partially, 13-28 completely

Isolated nucleic acid molecules encoding human prostate carcinom associated proteins as characterized by SEQIDs5,45,11,3 and SEQIDs 6,46,12,4, respectively; the recombinant expression of the same in host cells; the isolated proteins as characterized by SEQIDs 6,46,12,4; antisense RNA sequence and ribozyme complementary to said nucleic acid molecules; inhibitor that can suppress the activity of said prostate carcinom associated proteins; method for diagnosing a prostate carcinoma by contacting a sample with a nucleic acid, an antibody or other reagent that reacts with the mRNA of SEQIDs5,45,11,3; method for diagnosing endomertial cancer by contacting a target sample with a nucleic acid, an antibody or other reagent that reacts with the mRNA of SEQIDs5,45,11,3; method for diagnosing a melanoma, chorion carcinoma, cancer of the lung and ot the prostate comprising contacting a target sample with a reagent which detects antisense RNA of SEQIDs 11 and 3; method for preventing prostate tumour, endometrial cancer, choroin carcinoma or cancer of the lung comprising administering an inhibiting reagent of human prostate carcinom associated proteins; diagnostic kit containing an antibody; method for identifyng an agonist or an antagonist of human prostate carcinom associated proteins.

2. Claims: 1-12, 29-31 partially

Isolated nucleic acid molecule encoding human prostate carcinom associated protein as characterized by SEQIDs 7 and SEQIDs 8, respectively; the recombinant expression of the same in host cells; the isolated protein as characterized by SEQIDs 8; antisense RNA sequence and ribozyme complementary to said nucleic acid molecule; inhibitor that can suppress the activity of said prostate carcinom associated protein; diagnostic kit containing an antibody; method for identifyng an agonist or an antagonist of human prostate carcinom associated proteins.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 12 partially

Present claim 12 relates to an inhibitor wich is defined by reference to a desirable characteristic or property, namely suppressing the activity of the protein of claim 6.

The claims cover all inhibitors having this characteristic or property, whereas the application provides only support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for a limited number of such inhibitors.

In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the inhibitors by reference to a result to be achieved.
Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible.

Consequently, the search has been carried out for those parts of the claim 12 which appear to be clear, supported and disclosed, namely those parts relating to the Trp8/10 corresponding antibody , Trp8/10 corresponding antisense construct, a Trp8/10 corresponding ribozyme.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/EP 01/08309

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9909166 A	25-02-1999	US 6194152 B1	27-02-2001
		AU 744875 B2	07-03-2002
		AU 9021898 A	08-03-1999
		CA 2300364 A1	25-02-1999
		EP 1005549 A2	07-06-2000
		JP 2001514889 T	18-09-2001
		NZ 503404 A	01-03-2002
		WO 9909166 A2	25-02-1999
WO 0040614 A	13-07-2000	AU 2055600 A	24-07-2000
		CA 2360396 A1	13-07-2000
		EP 1141017 A2	10-10-2001
		JP 2002536966 T	05-11-2002
		WO 0040614 A2	13-07-2000
WO 9815657 A	16-04-1998	US 5919638 A	06-07-1999
		EP 0954599 A1	10-11-1999
		JP 2001523948 T	27-11-2001
		WO 9815657 A1	16-04-1998
		US 6110675 A	29-08-2000
WO 9837093 A	27-08-1998	US 6261562 B1	17-07-2001
		AU 731840 B2	05-04-2001
		AU 6181898 A	09-09-1998
		BR 9808881 A	11-09-2001
		CN 1252837 T	10-05-2000
		CZ 9903016 A3	13-03-2002
		EP 1005546 A2	07-06-2000
		HU 0002095 A2	28-10-2000
		NO 994069 A	22-10-1999
		NZ 337446 A	23-02-2001
		PL 335348 A1	25-04-2000
		TR 9902053 T2	21-04-2000
		US 6262245 B1	17-07-2001
		WO 9837093 A2	27-08-1998
		US 2002090372 A1	11-07-2002
		US 6465611 B1	15-10-2002
		US 6395278 B1	28-05-2002
		US 6329505 B1	11-12-2001
		US 2002022248 A1	21-02-2002
		US 2002051977 A1	02-05-2002
		US 2002193296 A1	19-12-2002
		ZA 9801585 A	04-09-1998
WO 0114423 A	01-03-2001	WO 0114423 A1	01-03-2001
WO 0104303 A	18-01-2001	AU 5778600 A	30-01-2001
		EP 1194546 A1	10-04-2002
		WO 0104303 A1	18-01-2001
WO 0142467 A	14-06-2001	AU 2074201 A	18-06-2001
		WO 0142467 A2	14-06-2001
WO 0151633 A	19-07-2001	AU 3447401 A	24-07-2001
		AU 6158700 A	30-01-2001
		WO 0104143 A2	18-01-2001
		EP 1194571 A1	10-04-2002
		EP 1261708 A2	04-12-2002

INTERNATIONAL SEARCH REPORT

Information on patent family members

mational Application No
PCT/EP 01/08309

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0151633 A		NO 20023402 A	29-08-2002
		WO 0151633 A2	19-07-2001
		US 2002022248 A1	21-02-2002
		US 2002051977 A1	02-05-2002
		US 2002193296 A1	19-12-2002
WO 0214361 A	21-02-2002	AU 8501801 A	25-02-2002
		WO 0214361 A2	21-02-2002
WO 0200722 A	03-01-2002	AU 7024001 A	08-01-2002
		WO 0200722 A2	03-01-2002
		US 2002156253 A1	24-10-2002
WO 0168857 A	20-09-2001	AU 4746001 A	24-09-2001
		WO 0168857 A2	20-09-2001
WO 0153348 A	26-07-2001	AU 3648201 A	31-07-2001
		EP 1252189 A2	30-10-2002
		WO 0153348 A2	26-07-2001
		US 2002072101 A1	13-06-2002
WO 0162794 A	30-08-2001	AU 3859601 A	03-09-2001
		WO 0162794 A2	30-08-2001
		US 2002142377 A1	03-10-2002
WO 0230268 A	18-04-2002	US 2002068036 A1	06-06-2002
		AU 1534502 A	22-04-2002
		WO 0230268 A2	18-04-2002

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 February 2002 (07.02.2002)

PCT

(10) International Publication Number
WO 02/010382 A2

(51) International Patent Classification[7]: C12N 15/12, 15/11, 9/00, C07K 14/47, C12Q 1/68, G01N 33/577, A61K 31/713

(21) International Application Number: PCT/EP01/08309

(22) International Filing Date: 18 July 2001 (18.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/221,513 28 July 2000 (28.07.2000) US

(71) Applicant and
(72) Inventor: WISSENBACH, Ulrich [DE/DE]; Institut für Pharmakologie und Toxikologie der Uni, versität des Saarlandes, 66421 Homburg (DE).

(74) Agent: HUBER, Bernard; Huber & Schüssler, Truderinger Str. 246, 81825 München (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

(48) Date of publication of this corrected version:
19 September 2002

(15) Information about Correction:
see PCT Gazette No. 38/2002 of 19 September 2002, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/010382 A2

(54) Title: TRP8, TRP9 AND TRP10, NOVEL MARKERS FOR CANCER

(57) Abstract: The present invention relates to gene expression in normal cells and cells of malignant tumors and particularly to novel markers associated with cancer, Trp8, Trp9 and Trp10, and the genes encoding Trp8, Trp9 and Trp10. Also provided are vectors, host cells, antibodies, and recombinant methods for producing these human proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating a tumor.

Trp8, Trp9 and Trp10, novel markers for cancer

FIELD OF THE INVENTION

The present invention relates to gene expression in normal cells and cells of malignant tumors and particularly to novel markers associated with cancer, Trp8, Trp9 and Trp10, and the genes encoding Trp8, Trp9 and Trp10

BACKROUND OF THE TECHNOLOGY

Prostate cancer is one of the most common diseases of older men world wide. Diagnosis and monitoring of prostate cancer is difficult because of the heterogeneity of the disease. For diagnosis different grades of malignancy can be distinguished according to the Gleason-Score Diagnosis. For this diagnosis a prostate tissue sample is taken from the patient by biopsy and the morphology of the tissue is investigated. However, this approach only yields subjective results depending on the experience of the pathologist. For confirmation of these results and for obtaining an early diagnosis an additional diagnostic method can be applied which is based on the detection of a prostate specific antigen (PSA). PSA is assayed in serum samples, blood samples etc. using an anti-PSA-antibody. However, since in principle PSA is also expressed in normal prostate tissue there is a requirement for the definition of a threshold value (about 4 ng/ml PSA) in order to be able to distinguish between normal and malign prostate tissue. Unfortunately, this diagnostic method is quite insensitive and often yields false-positive results. Moreover, by using this diagnostic method any conclusions as regards the grade of malignancy, the progression of the tumor and its potential for metastasizing cannot be drawn. Thus, the use of molecular markers would be helpful to distinguish benign from malign tissue and for grading and staging prostate carcinoma, particularly for patients with metastasizing prostate cancer having a very bad prognosis.

The above discussed limitations and failings of the prior art to provide meaningful specific markers which correlate with the presence of prostate tumors, in particular metastasizing tumors, has created a need for markers which can be used diagnostically, prognostically and therapeutically over the course of this disease. The present invention fulfils such a need by the provision of Tpr8, Trp9 and Trp10 and the genes encoding Trp8, Trp9 and Trp10: The genes encoding Trp8 and Trp10 are expressed in prostate carcinoma and prostatic metastasis, but

not in normal prostate, benign hyperplasia (BHP) and intraepithelial prostatic neoplasia (PIN). Furthermore, expression of Trp10 transcripts is detectable in carcinoma but not in healthy tissue of the lung, the prostate, the placenta and in melanoma.

SUMMARY OF THE INVENTION

The present invention is based on the isolation of genes encoding novel markers associated witha cancer, Trp8, Trp9 and Trp10. The new calcium channel proteins Trp8, Trp9 and Trp10 are members of the trp (transient receptor potential) - family, isolated from human placenta (Trp8a and Trp8b) and humane prostate (Trp9, Trp10a and Trp10b). Trp proteins belong to a steadily growing family of Ca^{2+} selective and non selective ion channels. In the recent years seven Trp proteins (trp1 - trp7) have been identified and suggested to be involved in cation entry, receptor operated calcium entry and pheromone sensory signaling. Structurally related to the trp proteins are the vanilloid receptor (VR1) and the vanilloid like receptor (VRL-1) both involved in nociception triggered by heat. Furthermore, two calcium permeable channels were identified in rat small intestine (CaT1) and rabbit kidney (ECaC). These distantly related channels are suggested to be involved in the uptake of calcium ions from the lumen of the small intestine (CaT1) or in the reuptake of calcium ions in the distal tubule of the kidney (ECaC). Common features or the Trp and related channels are a proposed structure comprising six transmembrane domains including several conserved amino acid motifs. In the present invention the cloning and expression of a CaT1 like calcium channel (Trp8) from human placenta as well as Trp9 and Trp10 (two variants, Trp10a and Trp10b) is described. Two polymorphic variants of the Trp8 cDNA were isolated from placenta (Trp8a and Trp8b). Transient expression of the Trp8b cDNA in HEK (human embryonic kidney) cells results in cytosolic calcium overload implicating that the Trp8 channel is constitutive open in the expression system. Trp8 induces highly calcium selective inward currents in HEK cells. The C -terminus of the Trp8 protein binds calmodulin in a calcium dependent manner. The Trp9 channel is expressed in trophoblasts and syncytiotrophoblasts of placenta and in pancreatic acinar cells. Furthermore, the Trp8 channel is expressed in prostatic carcinoma and prostatic metastases, but not in normal tissue of the prostate. No expression of Trp8 transcripts is detectable in benign prostatic hyperplasia (BPH) or prostatic intraepithelial neoplasia (PIN). Therefore, the Trp8 channel is exclusively expressed in malign prostatic tissues and serves as molecular marker for prostate cancer. From the experimental results it is also apparent that the

modulation of Trp8 and/or Trp10, e.g. the inhibition of expression or activity , is of therapeutic interest, e.g. for the prevention of tumor progression.

The present invention, thus, provides a Trp8, Trp9 and Trp10 protein, respectively, as well as nucleic acid molecule encoding the protein and, moreover, an antisense RNA, a ribozyme and an inhibitor, which allow to inhibit the expression or the activity of Trp8, Trp9 and/or Trp10.

In one embodiment, the present invention provides a diagnostic method for detecting a prostate cancer or endometrial cancer (cancer of the uterus) associated with Trp8 or Trp10 in a tissue of a subject, comprising contacting a sample containing Trp8 and/or Trp10 encoding mRNA with a reagent which detects Trp8 and/or Trp10 or the corresponding mRNA.

In a further embodiment, the present invention provides a diagnostic method for detecting a melanoma, chorion carcinoma, cancer of the lung and of the prostate in a tissue of a subject, comprising contacting a sample with a reagent which detects Trp10a and/or Trp10b antisense transcripts or Trp10a and/or Trp10b related antisense transcripts.

In another embodiment, the present invention provides a method of treating a prostate tumor, carcinoma of the lung, carcinoma of the placenta (chorion carcinoma) or melanoma associated with Trp8 and/or Trp10, comprising administering to a subject with such an disorder a therapeutically effect amount of a reagent which modulates, e.g. inhibits, expression of Trp8 and/or Trp10 or the activity of the protein, e.g. the above described compounds.

Finally, the present invention provides a method of gene therapy comprising introducing into cells of a subject an expression vector comprising a nucleotide sequence encoding the above mentioned antisense RNA or ribozyme, in operable linkage with a promoter.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1: A, phylogenetic relationship of trp and related proteins. B, hydropathy plot of the Trp8 protein sequence according to Kyte and Doolittle. C, alignment of Trp8a/b to the epithelial calcium channels ECaC (from rabbit) and Vr1 (from rat). Putative transmembrane domains are underlined.

Figure 2: A, polymorphism of the Trp8 gene. The polymorphic variants Trp8a and Trp8b differ in five base pairs resulting in three amino acid exchanges in the derived protein sequences. Specific primers were derived from the Trp8 gene as indicated by arrows. B, the Trp8a and Trp8b genes are distinguishable by a single restriction site. Genomic fragments of the Trp8 gene can be amplified using specific primers (shown in A). The genomic fragment of the Trp8b gene contains an additional site of the restriction enzyme BSP1286I (B). C, the Trp8 gene is located on chromosome 7. D, genotyping of eleven human subjects. A 458 bp genomic fragment of the Trp8 gene was amplified using specific primers (shown in A) and restricted with BSP1286I. The resulting fragments were analyzed by PAGE electrophoresis.

Figure 3: The Trp8b protein is a calcium selective ion channel. A, representative trace of a pdiTrp8b transfected HEK 293 cell. Trp8b mediated currents are activated by voltage ramps (-100 mV - +100 mV) of 100 msec at -40 mV or +70 mV holding potential. 1, Trp8b currents in the presence at 2mm $[Ca^{2+}]_o$;2, effect of solution switch alone 3, switch to nominal zero calcium solution. B, Trp8b currents in the presence of zero divalent cations. C, current voltage relationship of the currents shown in A. Inset, leak subtracted current. D, current voltage relationship of the current shown in B. E, statistics of representative experiments. Black: Trp8 transfected cells, gray: control cells. Columns from left to right: Trp8 currents at - 40 mV (n =12) and + 70 mV holding potential (n = 12). Trp8 currents in standard bath solution including 120 mM NMDG without sodium (n = 7) and with nominal zero calcium ions (n = 8) or in the presence of 1mM EGTA with zero divalent cations (n = 6). F, representative changes in $[Ca^{2+}]_i$ in Trp8b transfected HEK cells (gray) and controls (black) in the presence or absence of 1mM $[Ca^{2+}]_o$. Inset, relative increase of cytosolic calcium concentration of Trp8b transfected HEK cells, before and after readdition of 1 mM $[Ca^{2+}]_o$ in comparison to control cells.

Figure 4: The C-terminal region of the Trp8 protein binds calmodulin. A, N- and C-terminal fragments of the Trp8 protein used for calmodulin binding studies. B, the Trp8 protein and a truncated Trp8 protein which was in vitro translated after MunI cut of the cDNA, which lacks the C-terminal 32 amino acid residues, were in vitro translated in the presence of ^{35}S-methionine and incubated with calmodulin coupled agarose beads in the presence of 1 mM Ca^{2+} or 2 mM EGTA. C, calmodulin binding to N- and C-terminal fragments of the Trp8protein in the presence of Ca^{2+} (1 mM) or EGTA (2 mM)

Figure 5: Expression pattern of the Trp8 cDNA. A, Northern blots (left panels, Clontech, Palo Alto) were hybridized using a 348 bp NcoI/BamHT fragment of the Trp9 cDNA. The probe hybridizes to mRNA species isolated from the commercial blot, but not to mRNA species isolated from benign prostate hyperplasia (right panel, mRNA isolated from 20 human subjects with benign prostate hyperplasia). B,C, in situ hybridization with biotinylated Trp8 specific oligonucleotides on slides of human tissues. Left column antisense probes, right column sense probes. D, antinsense probes.

Figure 6: Differential expression of Trp8 cDNA in human prostate. A - F, in situ hybridization with prostatic tissues. A, normal prostate, B, primary carcinoma, C, benign hyperplasia, D, rezidive carcinoma, E, prostatic intraepithelial neoplasia, F, lymphnode metastasis of the prostata.

Figure 7: Trp8a cDNA sequence and derived amino acid sequence

Figure 8: A, Trp8b cDNA sequence and derived amino acid sequence
B, cDNA sequence of splice variant 1 (12B1)
C, cDNA sequence of splice variant 2 (17-3)
D, cDNA sequence of splice variant 3 (23A3)
E, cDNA sequence of splice variant 4 (23C3)

Figure 9: A, Trp9 cDNA sequence and derived amino acid sequence B, cDNA sequence of splice variant 15 and derived amino acid sequence.

Figure 10: A, cDNA sequence of Trp10a and derived amino acid sequence, B, cDNA fragment of Trp10a and derived amino acid sequence.

Figure 11: cDNA sequence of Trp10b and derived amino acid sequence.
Figure 12: Expression of Trp8 mRNA in human endometrial cancer or cancer of the uterus. A - D, in situ hybridization with slides of endometrial cancer hybridized with Trp8 antisense (left column) or sense probes as controls (right column). E - F, Trp8 antisense probes hybridized to slides of normal endometrium. It can be clearly seen no hybridization occurs with normal endometrial tissue.

Figure 13: Expression of human Trp9 and Trp10 genes
Northern blots were hybridized using Trp9 (upper panel) or Trp10 (lower panel) specific probes. Expression of the Trp9 cDNA is detectable in many tissues including human prostate and colon as well as in benign prostatic hyperplasia. Expression of Trp10 cDNA is detectable in human prostate of a commercial northern blot (Clontech, right side). This Northern blot contains prostatic tissue collected from 15 human subjects in the range of 14 - 60 years of age. No expression of Trp10 cDNA was detectable in benign prostatic hyperplasia (left side).

Figure 14: Expression of Trp10 transcripts and Trp10-antisense transcripts in human prostate cancer and metastasis of a melanoma. In situ hybridizations of slides hybridized with Trp10-antisense (A-E, K-N) and Trp10 related sense probes (F-J, P-R). It can clearly be seen that both probes detect the same cancer cells indicating that these cancer cells express Trp10 transcripts as well as Trp10-antisense transcripts. S, no Trp10 expression is detectable in benign hyperplasia of the prostate (BPH). O and T, show expression of Trp10 transcripts (O) and Trp10-antisense transcripts (T) in a metastasis of a melanoma in human lung. Melanoma cancer cells express both Trp10 transcripts and Trp10-antisense transcripts.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to an isolated nucleic acid molecule encoding the human prostate carcinoma associated protein Trp8a, Trp8b, Trp9, Trp10a or Trp10b or a protein exhibiting biological properties of Trp8a, Trp8b, Trp9, Trp10a or Trp10b and being selected from the group consisting of

(a) a nucleic acid molecule encoding a protein that comprises the amino acid sequence depicted in Figure 7, 8A, 9,10 or 11;

(b) a nucleic acid molecule comprising the nucleotide sequence depicted in Figure 7, 8A, 9,10, or 11;

(c) a nucleic acid molecule included in DSMZ Deposit no. DSM 13579 (deposit date: 28 June 2000), DSM 13580 (deposit date: 28 June 2000), DSM 13584 (deposit date: 5 July 2000), DSM 13581 (deposit date: 28 June 2000) or DSM ....(deposit date:....);

(d) a nucleic acid molecule with hybridizes to a nucleic acid molecule specified in (a) to (c)

(e) a nucleic acid molecule the nucleic acid sequence of which deviates from the nucleic sequences specified in (a) to (d) due to the degeneration of the genetic code; and

(f) a nucleic acid molecule, which represents a fragment, derivative or allelic variation of a nucleic acid sequence specified in (a) to (e).

As used herein, a protein exhibiting biological properties of Trp8a, Trp8b, Trp9,Trp10a or Trp10b is understood to be a protein having at least one of the activities as illustrated in the Examples, below.

As used herein, the term „isolated nucleic acid molecule„ includes nucleic acid molecules substantially free of other nucleic acids, proteins, lipids, carbohydrates or other materials with which it is naturally associated.

In a first embodiment, the invention provides an isolated nucleic acid molecule encoding the human prostate carcinoma associated protein Trp8a, Trp8b, Trp9,Trp10a or Trp10b comprising the amino acid sequence depicted in Figure 7, 8A, 9,10 or 11. The present invention also provides a nucleic acid molecule comprising the nucleotide sequence depicted in Figure 7, 8A, 9,10 or 11.

The present invention provides not only the generated nucleotide sequence identified in Figure 7, 8A, 9,10 or 11, respectively and the predicted translated amino acid sequence, respectively, but also plasmid DNA containing a Trp8a cDNA deposited with the DSMZ, under DSM 13579, a Trp8b cDNA deposited with the DSMZ, under DSM 13580, a Trp9 cDNA deposited with the DSMZ, under DSM 13584, a Trp10a cDNA deposited with the DSMZ, under DSM 13581, and a Trp10b cDNA deposited with the DSMZ, under DSM...., respectively. The nucleotide sequence of each deposited Trp-clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by each deposited clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited Trp-encoding DNA, collecting the protein, and determining its sequence.

The nucleic acid molecules of the invention can be both DNA and RNA molecules. Suitable DNA molecules are, for example, genomic or cDNA molecules. It is understood that all

nucleic acid molecules encoding all or a portion of Trp8a, Trp8b, Trp9,Trp10a or Trp10b are also included, as long as they encode a polypeptide with biological activity. The nucleic acid molecules of the invention an be isolated from natural sources or can be synthesized according to know methods.

The present invention also provides nucleic acid molecules which hybridize to the above nucleic acid molecules. As used herein, the term „hybridize„ has the meaning of hybridization under conventional hybridization conditions, preferably under stringent conditions as described, for example, in Sambrook et al., Molecular Cloning, A Laboratory Manual 2nd edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. Also contemplated are nucleic acid molecules that hybridize to the Trp nucleic acid molecules at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency), salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°Cin a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 9.2M NaH_2PO_4; 0.02M EDTA, pH7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA, following by washes at 50°C with 1 X SSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC). Variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Nucleic acid molecules that hybridize to the molecules of the invention can be isolated, e.g., from genomic or cDNA libraries that were produced from human cell lines or tissues. In order to identify and isolate such nucleic acid molecules the molecules of the invention or parts of these molecules or the reverse complements of these molecules can be used, for example by means of hybridization according to conventional methods (see, e.g., Sambrook et al., supra). As a hybridization probe nucleic acid molecules can be used, for example, that have exactly or basically the nucleotide sequence depicted in Figure 7, 8A, 9,10 or 11, respectively, or parts of these sequences. The fragments used as hybridization probe can be synthetic

fragments that were produced by means of conventional synthetic methods and the sequence of which basically corresponds to the sequence of a nucleic acid molecule of the invention.

The nucleic acid molecules of the present invention also include molecules with sequences that are degenerate as a result of the genetic code.

In a further embodiment, the present invention provides nucleic acid molecules which comprise fragments, derivatives and allelic variants of the nucleic acid molecules described above encoding a protein of the invention. „Fragments„ are understood to be parts of the nucleic acid molecules that are long enough to encode one of the described proteins. These fragments comprise nucleic acid molecules specifically hybridizing to transcripts of the nucleic acid molecules of the invention. These nucleic acid molecules can be used, for example, as probes or primers in the diagnostic assay and/or kit described below and, preferably, are oligonucleotides having a length of at least 10, in particular of at least 15 and particularly preferred of at least 50 nucleotides. The nucleic acid molecules and oligonucleotides of the invention can also be used, for example, as primers for a PCR reaction. Examples of particular useful probes (primers) are shown in Tables 1 and 2.

Table 1

Trp8 probes used for in situ hybridization:

Probes (antisense)

1.) 5' TCCGCTGCCGGTTGAGATCTTGCC 3'

2.) 5' CTTGCTCCATAGGCAGAGAATTAG 3'

3.) 5' ATCCTCAGAGCCCCGGGTGTGGAA3'

Controls (sense)

1.) 5' GGCAAGATCTCAACCGGCAGCGGA 3'

2.) 5' CTAATTCTCTGCCTATGGAGCAAG 3'

3.) 5' TTCCACACCCGGGGCTCTGAGGAT 3'

Tabelle 2

Trp10 probes used for the in situ hybridizations shown in Figure 14:

Probes (antisense)

1.) 5' GCTTCCACCCCAAGCTTCACAGGAATAGA 3' (Figure 14 A, 14B)

2.) 5' GGCGATGAAATGCTGGTCTGTGGC 3' (Figure 14C, 14D, 14N, 14S, 14O)

3.) 5' ATCTTCCAGTTCTTGGTGTCTCGG 3' (Figure 14E, 14K)

4.) 5' GCTGCAGTACTCCTGCACCAGGAA 3' (Figure 14L, 14M)

Probes (sense)

1.) 5' TCTATTCCTGTGAAGCTTGGGGTGGAAGC 3' (Figure 14F, 14G)

2.) 5' GCCACAGACCAGCATTTCATCGCC 3' (Figure 14H, 14I, 14T)

3.) 5' CCGAGACACCAAGAACTGGAAGAT 3' (Figure 14J, 14P)

4.) 5' TTCCTGGTGCAGGAGTACTGCAGC 3' (Figure 14Q, 14R)

The term „derivative„ in this context means that the sequences of these molecules differ from the sequences of the nucleic acid molecules described above at one or several positions but have a high level of homology to these sequences. Homology hereby means a sequence identity of at least 40%, in particular an identity of at least 60%, preferably of more than 80% and particularly preferred of more than 90%. These proteins encoded by the nucleic acid molecules have a sequence identity to the amino acid sequence depicted in Figure 7, 8A, 9, 10 and 11, respectively, of at least 80%, preferably of 85% and particularly preferred of more than 90%, 97% and 99%. The deviations to the above-described nucleic acid molecules may have been produced by deletion, substitution, insertion or recombination. The definition of the derivatives also includes splice variants, e.g. the splice variants shown in Figures 8B to 8E and 9B.

The nucleic acid molecules that are homologous to the above-described molecules and that represent derivatives of these molecules usually are variations of these molecules that represent modifications having the same biological function. They can be naturally occurring variations, for example sequences from other organisms, or mutations that can either occur naturally or that have been introduced by specific mutagenesis. Furthermore the variations can be synthetically produced sequences. The allelic variants can be either naturally occurring variants or synthetically produced variants or variants produced by recombinant DNA processes.

Generally, by means of conventional molecular biological processes it is possible (see, e.g., Sambrook et al., supra) to introduce different mutations into the nucleic acid molecules of the invention. As a result Trp proteins or Trp related proteins with possibly modified biological properties are synthesized. One possibility is the production of deletion mutants in which nucleic acid molecules are produced by continuous deletions from the 5'- or 3'-terminal of the coding DNA sequence and that lead to the synthesis of proteins that are shortened accordingly. Another possibility is the introduction of single-point mutation at positions where a modification of the amino aid sequence influences, e.g., the ion channel properties or the regulations of the trp-ion channel. By this method muteins can be produced, for example, that possess a modified ion conducting pore, a modified K_m-value or that are no longer subject to the regulation mechanisms that normally exist in the cell, e.g. with regard to allosteric regulation or covalent modification. Such muteins might also be valuable as therapeutically useful antagonists of Trp8a, Trp8b, Trp9,Trp10a or Trp10b, respectively.

For the manipulation in prokaryotic cells by means of genetic engineering the nucleic acid molecules of the invention or parts of these molecules can be introduced into plasmids allowing a mutagenesis or a modification of a sequence by recombination of DNA sequences. By means of conventional methods (cf. Sambrook et al., supra) bases can be exchanged and natural or synthetic sequences can be added. In order to link the DNA fragments with each other adapters or linkers can be added to the fragments. Furthermore, manipulations can be performed that provide suitable cleavage sites or that remove superfluous DNA or cleavage sites. If insertions, deletions or substitutions are possible, in vitro mutagenesis, primer repair, restriction or ligation can be performed. As analysis method usually sequence analysis, restriction analysis and other biochemical or molecular biological methods are used.

The proteins encoded by the various variants of the nucleic acid molecules of the invention show certain common characteristics, such as ion channel activity, molecular weight, immunological reactivity or conformation or physical properties like the electrophoretical mobilty, chromatographic behavior, sedimentation coefficients, solubility, spectroscopic properties, stability; pH optimum, temperature optimum.

The invention furthermore relates to vectors containing the nucleic acid molecules of the invention. Preferably, they are plasmids, cosmids, viruses, bacteriophages and other vectors

usually used in the field of genetic engineering. Vectors suitable for use in the present invention include, but are not limited to the T7-based expression vector for expression in mammalian cells and baculovirus-derived vectors for expression in insect cells. Preferably, the nucleic acid molecule of the invention is operatively linked to the regulatory elements in the recombinant vector of the invention that guarantee the transcription and synthesis of an RNA in prokryotic and/or eukaryotic cells that can be translated. The nucleotide sequence to be transcribed can be operably linked to a promoter like a T7, metallothionein I or polyhedrin promoter.

In a further embodiment, the present invention relates to recombinant host cells transiently or stable containing the nucleic acid molecules or vectors or the invention. A host cell is understood to be an organism that is capable to take up *in vitro* recombinant DNA and, if the case may be, to synthesize the proteins encoded by the nucleic acid molecules of the invention. Preferably, these cells are prokaryotic or eukaryotic cells, for example mammalian cells, bacterial cells, insect cells or yeast cells. The host cells of the invention are preferably characterized by the fact that the introduced nucleic acid molecule of the invention either is heterologous with regard to the transformed cell, i.e. that it does not naturally occur in these cells, or is localized at a place in the genome different from that of the corresponding naturally occurring sequence.

A further embodiment of the invention relates to isolated proteins exhibiting biological properties of the human prostate carcinoma associated protein Trp8a, Trp8b, Trp9,Trp10a or Trp10b and being encoded by the nucleic acid molecules of the invention, as well as to methods for their production, whereby, e.g., a host cell of the invention is cultivated under conditions allowing the synthesis of the protein and the protein is subsequently isolated from the cultivated cells and/or the culture medium. Isolation and purification of the recombinantly produced proteins may be carried out by conventional means including preparative chromatography and affinity and immunological separations involving affinity with an anti-Trp8a-, anti-Trp8b-, anti-Trp9-,anti-Trp10a- or anti-Trp10b-antibody, respectively.

As used herein, the term „isolated protein„ includes proteins substantially free of other proteins, nucleic acids, lipids, carbohydrates or other materials with which it is naturally associated. Such proteins however not only comprise recombinantly produced proteins but include isolated naturally occurring proteins, synthetically produced proteins, or proteins

produced by a combination of these methods. Means for preparing such proteins are well understood in the art. The Trp proteins are preferably in a substantially purified form. A recombinantly produced version of a human prostate carcinoma associated protein Trp8a, Trp8b, Trp9,Trp10a or Trp10b protein, including the secreted protein, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67; 31-40 (1988).

In a further preferred embodiment, the present invention relates to an antisense RNA sequence characterised that it is complementary to an mRNA transcribed from a nucleic acid molecule of the present invention or a part thereof and can selectively bind to said mRNA, said sequence being capable of inhibiting the synthesis of the protein encoded by said nucleic acid molecules, and a ribozyme characterised in that it is complementary to an mRNA transcribed from a nucleic acid molecule of the present invention or a part thereof and can selectively bind to and cleave said mRNA, thus inhibiting the synthesis of the proteins encoded by said nucleic acid molecules. Riboyzmes which are composed of a single RNA chain are RNA enzymes, i.e. catalytic RNAs, which can intermolecularly cleave a target RNA, for example the mRNA transcribed from one of the Trp genes. It is now possible to construct ribozymes which are able to cleave the target RNA at a specific site by following the strategies described in the literature. (see, e.g., Tanner et al., in: Antisense Research and Applications, CRC Press Inc. (1993), 415-426). The two main requirements for such ribozymes are the catalytic domain and regions which are complementary to the target RNA and which allow them to bind to its substrate, which is a prerequisite for cleavage. Said complementary sequences, i.e., the antisense RNA or ribozyme, are useful for repression of Trp8a-, Trp8b, Trp9-,Trp10a- and Trp10b-expression, respectively, i.e. in the case of the treatment of a prostate cancer or endometrial cancer (carcinoma of the uterus). Preferably, the antisense RNA and ribozyme of the invention are complementary to the coding region. The person skilled in the art provided with the sequences of the nucleic acid molecules of the present invention will be in a position to produce and utilise the above described antisense RNAs or ribozymes. The region of the antisense RNA and ribozyme, respectively, which shows complementarity to the mRNA transcribed from the nucleic acid molecules of the present invention preferably has a length of at least 10, in particular of at least 15 and particularly preferred of at least 50 nucleotides.

In still a further embodiment, the present invention relates to inhibitors of Trp8a, Trp8b, Trp9, Trp10a and Trp10b, respectively, which fulfill a similar purpose as the antisense RNAs or

ribozymes mentioned above, i.e. reduction or elimination of biologically active Trp8a, Trp8b, Trp9, Trp10a or Trp10b molecules. Such inhibitors can be, for instance, structural analogues of the corresponding protein that act as antagonists. In addition, such inhibitors comprise molecules identified by the use of the recombinantly produced proteins, e.g. the recombinantly produces protein can be used to screen for and identify inhibitors, for example, by exploiting the capability of potential inhibitors to bind to the protein under appropriate conditions. The inhibitors can, for example, be identified by preparing a test mixture wherein the inhibitor candidate is incubated with Trp8a, Trp8b, Trp9, Trp10a or Trp10b, respectively, under appropriate conditions that allow Trp8a, Trp8b, Trp9, Trp10a or Trp10b to be in a native conformation. Such an in vitro test system can be established according to methods well known in the art. Inhibitors can be identified, for example, by first screening for either synthetic or naturally occurring molecules that bind to the recombinantly produced Trp protein and then, in a second step, by testing those selected molecules in cellular assays for inhibition of the Trp protein, as reflected by inhibition of at least one of the biological activities as described in the examples, below. Such screening for molecules that bind Trp8a, Trp8b, Trp9, Trp10a or Trp10b could easily performed on a large scale, e.g. by screening candidate molecules from libraries of synthetic and/or natural molecules. Such an inhibitor is, e.g., a synthetic organic chemical, a natural fermentation product, a substance extracted from a microorganism, plant or animal, or a peptide. Additional examples of inhibitors are specific antibodies, preferably monoclonal antibodies. Moreover, the nucleic sequences of the invention and the encoded proteins can be used to identify further factors involved in tumor development and progression. In this context it should be emphasized that the modulation of the calcium channel of a member of the trp family can result in the stimulation of the immune response of T lymphocytes leading to proliferation of the T lymphocytes. The proteins of the invention can, e.g., be used to identify further (unrelated) proteins which are associated with the tumor using screening methods based on protein/protein interactions, e.g. the two-hybrid-system Fields, S. and Song, O. (1989) Nature (340): 245-246.

The present invention also provides a method for diagnosing a prostate carcinoma which comprises contacting a target sample suspected to contain the protein Trp8a, Trp8b, Trp10a and/or Trp10b or the Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA with a reagent which reacts with Trp8a, Trp8b, Trp10a and/or Trp10b or the Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA and detecting Trp8a, Trp8b, Trp10a and/or Trp10b or Trp8a, Trp8b, Trp10a and/or Trp 10b encoding mRNA.

It has been found that carcinoma cells of placenta (chorion carcinoma), lung and prostate express Trp10 transcripts as well as Trp10 antisense transcripts and transcripts being in part complementary to Trp10 antisense transcripts. Accordingly, the present invention also provides a method for diagnosing a melanoma, chorion carcinoma, cancer of the lung and of the prostate in a tissue of a subject, comprising contacting a sample with a reagent which detects Trp10a and/or Trp10b antisense RNA.

When the target is mRNA (or antisense RNA), the reagent is typically a nucleic acid probe or a primer for PCR. The person skilled in the art is in a position to design suitable nucleic acids probes based on the information as regards the nucleotide sequence of Trp8a, Trp8b, Trp10a or Trp10b as depicted in figure 7, 8a, 10 and 11, respectively, or tables 1 and 2, above. When the target is the protein, the reagent is typically an antibody probe. The term „antibody", preferably, relates to antibodies which consist essentially of pooled monoclonal antibodies with different epitopic specifities, as well as distinct monoclonal antibody preparations. Monoclonal antibodies are made from an antigen containing fragments of the proteins of the invention by methods well known to those skilled in the art (see, e.g., Köhler et al., Nature 256 (1975), 495). As used herein, the term „antibody" (Ab) or „monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab') 2 fragments) which are capable of specifically binding to protein. Fab and f(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24: 316-325 (1983)). Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimerical, single chain, and humanized antibodies. The target cellular component, i.e. Trp8a, Trp8b, Trp10a and/or Trp10b or Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA or Trp10a/b antisense transcripts, e.g., in biological fluids or tissues, may be detected directly in situ, e.g. by in situ hybridization (e.g., according to the examples, below) or it may be isolated from other cell components by common methods known to those skilled in the art before contacting with a probe. Detection methods include Northern blot analysis, RNase protection, in situ methods, e.g. in situ hybridization, in vitro amplification methods (PCR, LCR, QRNA replicase or RNA-transcription/amplification (TAS, 3SR), reverse dot blot disclosed in EP-B1 O 237 362)), immunoassays, Western blot and other detection assays that are known to those skilled in the art.

Products obtained by in vitro amplification can be detected according to established methods, e.g. by separating the products on agarose gels and by subsequent staining with ethidium bromide. Alternatively, the amplified products can be detected by using labeled primers for amplification or labeled dNTPs.

The probes can be detectable labeled, for example, with a radioisotope, a bioluminescent, compound, a chemiluminescent compound, a fluorescent compound, a metal chelate, or an enzyme.

Expression of Trp8a, Trp8b, Trp10a and Trp10b, respectively, in tissues can be studied with classical immunohistological methods (Jalkanen et al., J. Cell. Biol. 101 (1985), 976-985; Jalkanen et al., J. Cell. Biol. 105 (1987), 3087-3096; Sobol et al. Clin. Immunpathol. 24 (1982), 139-144; Sobol et al., Cancer 65 (1985), 2005-2010). Other antibody based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (^{125}I, ^{121}I), carbon (^{14}C), sulfur (^{35}S), tritium (^{3}H), indium (^{112}In), and technetium rhodamine, and biotin. In addition to assaying Trp8a, Trp8b, Trp 10a or Trp10b levels in a biological sample, the protein can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma. A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, ^{131}I, ^{112}In, ^{99}mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of ^{99}mTc. The labeled antibody or antibody fragment will then preferentially accumulate at he location of cells which contain the specific protein. In

vivo tumor imaging is described in S.W. Burchiel et al., „Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments". (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B.A. Rhodes, eds., Masson Publishing Inc. (1982)).

The marker Trp8a and Trp8b is also useful for prognosis, for monitoring the progression of the tumor and the diagnostic evaluation of the degree of malignancy of a prostate tumor (grading and staging), e.g. by using in situ hybridization: In a primary carcinoma Trp8 is expressed in about 2 to 10% of carcinoma cells, in a rezidive carcinoma in about 10 to 60% of cells and in metastases in about 60 to 90% of cells.

The present invention also relates to a method for diagnosing endometrial cancer (cancer of the uterus) which comprises contacting a target sample suspected to contain the protein Trp8a and/or Trp8b or the Trp8a and/or Trp8b encoding mRNA with a reagent which reacts with Trp8a and/or Trp8b or the encoding mRNA and detecting Trp8a and/or Trp8b encoding mRNA. As regards particular embodiments of this method reference is made to the particular embodiments of the method of diagnosing a prostate cancer outlined above.

For evaluating whether the concentration of Trp8a, Trp8b, Trp10a or Trp10b or the concentration of Trp8a, Trp8b, Trp10a or Trp10b encoding mRNA is normal or increased, thus indicative for the presence of a malignant tumor, the measured concentration is compared with the concentration in a normal tissue, preferably by using the ratio of Trp8a:Trp9, Trp8b:Trp9 or Trp10(a or b)/Trp9 for quantification.

Since the prostate carcinoma forms its own basement membrane when growing invasively, it can be concluded that only cells expressing Trp8 and Trp10 are involved in this phenomenon. Thus, it can be concluded that by inhibiting the expression and/or activity of these proteins an effective therapy of cancers like PCA is provided.

Thus, the present invention also relates to a pharmaceutical composition containing a reagent which decreases or inhibits Trp8a, Trp8b, Trp10a and/or Trp10b expression or the activity of Trp8a, Trp8b, Trp10a and/or Trp10b, and a method for preventing, treating, or ameliorating a prostate tumor, endometrial cancer (uterine carcinoma) tumor, a chorion carcinoma, cancer of the lung or melanoma,which comprises administering to a mammalian subject a

therapeutically effective amount of a reagent which decreases or inhibits Trp8a, Trp8b, Trp10a and/or Trp10b expression or the activity of Trp8a, Trp8b, Trp10a and/or Trp10b. Examples of such reagents are the above described antisense RNAs, ribozymes or inhibitors, e.g. specific antibodies. Furthermore, peptides, which inhibit or modulate the biological function of Trp8a, Trp8b, Trp9, Trp10a and/or Trp10b may be useful as therapeutical reagents. For example, these peptides can be obtained by screening combina torial phage display libraries (Cosmix, Braunschweig, Germany) as described by Rottgen, P. and Collins, J. (Gene (1995) 164 (2): 243-250). Furthermore, antigenic epitopes of the Trp8 and Trp10 proteins can be identified by the expression of recombinant Trp8 and Trp10 epitope libraries in E. coli (Marquart, A. & Flockerzi, V., FEBS Lett. 407 (1997), 137-140; Trost, C., et al., FEBS Lett. 451 (1999) 257-263 and the consecutive screening of these libraries with serum of patients with cancer of the prostate or of the endometrium. Those Trp8 and Trp10 epitopes which are immunogenic and which lead to the formation of antibodies in the serum of the patients can be then be used as Trp8 or Trp10 derived peptide vaccines for immune inventions against cancer cells which express Trp8 or Trp10. Alternatively to the E. coli expression system, Trp8 or Trp10 or epitopes of Trp8 and Trp10 can be expressed in mammalian cell lines such as human embryonic kidney (Hek 293) cells (American Type Culture Collection, ATCC CRL 1573).

Finally, compounds useful for therapy of the above described diseases comprise compounds which act as antagonists or agonists on the ion channels Trp8, Trp9 and Trp10. It could be shown that Trp8 is a highly calcium selective ion channel which in the presence of monovalent (namely sodium) and divalent ions (namely calcium) is only permeable for calcium ions (see Example 4, below, and Figures 3A, C, E). Under physiological conditions, Trp8 is a calcium selective channel exhibiting large inward currents. This very large conductance of Trp8 channels (as wells as Trp9 and Trp10a/b channels) is useful to establish systems for screening pharmacological compounds interacting with Trp-channels including high throughput screening systems. Useful high throughput screening systems are well known to the person skilled in the art and include, e.g., the use of cell lines stably or transiently transfected with DNA sequences encoding Trp8, Trp9 and Trp10 channels in assays to detect calcium signaling in biological systems. Such systems include assays based on Ca-sensitive dyes such as aequorin, apoaequorin, Fura-2, Fluo-3 and Indo-1.

Accordingly, the present invention also relates to a method for identifying compounds which act as agonists or antagonists on the ion channels Trp8, Trp9 and/or Trp10, said method comprising contacting a test compound with the ion channel Trp8, Trp9 and/or Trp10, preferably by using a system based on cells stably or transiently transfected with DNA sequences encoding Trp8, Trp9 and/or Trp10, and determining whether said test compound affects the calcium uptake.

For administration the above described reagents are preferably combined with suitable pharmaceutical carriers. Examples of suitable pharmaceutical carriers are well known in the art and include phosphate buffered saline solutions, water, emulsions, such as oil/water emulsions, various types of wetting agents, sterile solutions etc. Such carriers can be formulated by conventional methods and can be administered to the subject at a suitable dose. Administration of the suitable compositions may be effected by different ways, e.g. by intravenous, intraperetoneal, subcutaneous, intramuscular, topical or intradermal administration. The route of administration, of course, depends on the nature of the tumor and the kind of compound contained in the pharmaceutical composition. The dosage regimen will be determined by the attending physician and other clinical factors. As is well known in the medical arts, dosages for any one patient depends on many factors, including the patient's size, body surface area, age, sex, the particular compound to be administered, time and route of administration, the kind and stage of the tumor, general health and other drugs being administered concurrently.

The delivery of the antisense RNAs or ribozymes of the invention can be achieved by direct application or, preferably, by using a recombinant expression vector such as a chimeric virus containing these compounds or a colloidal dispersion system. By delivering these nucleic acids to the desired target, the intracellular expression of Trp8a, Trp8b, Trp10a and/or Trp10b and, thus, the level of Trp8a, Trp8b, Trp10a and/or Trp10b can be decreased resulting in the inhibition of the negative effects of Trp8a, Trp8b, Trp10a and/or Trp10b, e.g. as regards the metastasis formation of PCA.

Direct application to the target site can be performed, e.g., by ballistic delivery, as a colloidal dispersion system or by catheter to a site in artery. The colloidal dispersion systems which can be used for delivery of the above nucleic, acids include macromolecule complexes, nanocapsules, microspheres, beads and lipid-based systems including oil-in-water emulsions

(mixed), micelles, liposomes and lipoplexes, The preferred colloidal system is a liposome. The composition of the liposome is usually a combination of phospholipids and steroids, especially cholesterol. The skilled person is in a position to select such liposomes which are suitable for the delivery of the desired nucleic acid molecule. Organ-specific or cell-specific liposomes can be used in order to achieve delivery only to the desired tumor. The targeting of liposomes can be carried out by the person skilled in the art by applying commonly known methods. This targeting includes passive targeting (utilizing the natural tendency of the liposomes to distribute to cells of the RES in organs which contain sinusoidal capillaries) or active targeting (for example by coupling the liposome to a specific ligand, e.g., an antibody, a receptor, sugar, glycolipid, protein etc., by well known methods). In the present invention monoclonal antibodies are preferably used to target liposomes to specific tumors via specific cell-surface ligands.

Preferred recombinant vectors useful for gene therapy are viral vectors, e.g. adenovirus, herpes virus, vaccinia, or, more preferably, an RNA virus such as a Retrovirus. Even more preferably, the retroviral vector is a derivative of a murine or avian retrovirus. Examples of such retroviral vectors which can be used in the present invention are: Moloney murine leukemia virus (MoMuLV), Harvey murine sarcoma virus (HaMuSV), murine mammary tumor virus (MuMTV) and Rous sarcoma virus (RSV). Most preferably, a non-human primate retroviral vector is employed, such as the gibbon ape leukemia virus (GaLV), providing a broader host range compared to murine vectors. Since recombinant retroviruses are defective, assistance is required in order to produce infectious particles. Such assistance can be provided, e.g., by using helper cell lines that contain plasmids encoding all of the structural genes of the retrovirus under the control of regulatory sequences within the LTR. Suitable helper cell lines are well known to those skilled in the art. Said vectors can additionally contain a gene encoding a selectable marker so that the transduced cells can be identified. Moreover, the retroviral vectors can be modified in such a way that they become target specific. This can be achieved, e.g., by inserting a polynucleotide encoding a sugar, a glycolipid, or a protein, preferably an antibody. Those skilled in the art know additional methods for generating target specific vectors. Further suitable vectors and methods for in vitro- or in vivo-gene therapy are described in the literature and are known to the persons skilled in the art; see, e.g., WO 94/29469 or WO 97/00957.

In order to achieve expression only in the target organ, i.e. tumor to be treated, the nucleic acids encoding, e.g. an antisense RNA or ribozyme can also be operably linked to a tissue specific promoter and used for gene therapy. Such promoters are well known to those skilled in the art (see e.g. Zimmermann et al., (1994) Neuron 12, 11-24; Vidal et al.; (1990) EMBO J. 9, 833-840; Mayford et al., (1995), Cell 81, 891-904; Pinkert et al., (1987) Genes & Dev. 1, 268-76).

For use in the diagnostic research discussed above, kits are also provided by the present invention. Such kits are useful for the detection of a target cellular component, which is Trp8a, Trp8b, Trp10a and/or Trp10b or, alternatively, Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA or Trp10a/b antisense transcripts, wherein the presence or an increased concentration of Trp8a, Trp8b, Trp10a and/or Trp10b or, alternatively, Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA or Trp10a/b antisense transcripts is indicative for a prostate tumor, endometrial cancer, melanoma, chorion carcinoma or cancer of the lung, said kit comprising a probe for detection of Trp8a, Trp8b, Trp9, Trp10a and/or Trp10b or, alternatively, Trp8a, Trp8b, Trp9, Trp10a and/or Trp10b encoding mRNA or Trp10a/b antisense transcripts. The probe can be detectably labeled. Such probe may be a specific antibody or specific oligonucleotide. In a preferred embodiment, said kit contains an anti-Trp8a-, anti-Trp8b-, anti-Trp9-, anti-Trp10a-and/or anti-Trp10b-antibody and allows said diagnosis, e.g., by ELISA and contains the antibody bound to a solid support, for example, a polystyrene microtiter dish or nitrocellulose paper, using techniques known in the art. Alternatively, said kits are based on a RIA and contain said antibody marked with a radioactive isotope. In a preferred embodiment of the kit of the invention the antibody is labeled with enzymes, fluorescent compounds, luminescent compounds, ferromagnetic probes or radioactive compounds. The kit of the invention may comprise one or more containers filled with, for example, one or more probes of the invention. Associated with container (s) of the kit can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, us or sale for human administration.

EXAMPLES

The following Examples are intended to illustrate, but not to limit the invention. While such Examples are typical of those that might be used, other methods known to those skilled in the art may alternatively be utilized.

Example 1: Materials and Methods

(A) Isolation of cDNA clones and Northern blot analysis

Total RNA was isolated from human placenta an prostate using standard techniques. Isolation of mRNA was performed with poly $(A)^+$RNA - spin columns (New England Biolabs, Beverly, USA) according to the instructions of the manufacturer. Poly (a) $^+$RNA was reverse transcribed using the cDNA choice system (Gibco-BRL, Rockville, USA) and subcloned in λ-Zap phages (Stratagene, La Jolla, USA). An human expressed sequence tag (GenBank accession number 1404042) was used to screen an oligo d(T) primed human placenta cDNA library. Several cDNA clones were identified and isolated. Additional cDNA clones were isolated from two specifically primed cDNA libraries using primers 5'-gca tag gaa ggg aca ggt gg-3' and 5'-gag agt cga ggt cag tgg tcc-3'.

cDNA clones were sequenced using a thermocycler (PE Applied Biosystems, USA) and Thermo Sequenase (Amersham Pharmacia Biotech Europe, Freiburg, Germany). DNA sequences were analyzed with an automated sequencer (Licor, Linccoln, USA).

For Northern blot analysis 5 µg human poly $(A)^+$ RNA from human placenta or prostate were separated by electrophoresis on 0.8 % agarose gels. Poly $(A)^+$ RNA was transferred to Hybond N nylon membranes (Amersham Pharmacia Biotech Europe, Freiburg, Germany). The membranes were hyridized in the presence of 50 % formamide at 42^0C over night. DNA probes were labelled using [α^{32}P]dCTP and the „ready prime„ labelling kit (Amersham Pharmacia Biotech Europe, Freiburg, Germany). Commercial Northern blots were hybridized according to the distributors instructions (Clontech, Paolo Alto, USA).

(B) Construction of expression plasmids and transfection of HEK 293 cells

Lipofections were carried out with the recombinant dicistronic eucaryotic expression plasmid pdiTRP8 containing the cDNA of Trp8b under the control of the chicken ß-actin promotor followed by an internal ribosome entry side (IRES) and the cDNA of the green fluorescent protein (GFP). To obtain pdiTRP8 carrying the entire protein coding regions of TRP8b and

the GFP (Prasher, D.C. et al. (1992), Gene 111, 229-233), the 5`and 3´-untranslated sequences of the TRP8b cDNA were removed, the consensus sequence for initiation of translation in vertebrates (Kozak, M. (1987) Nucleic Acids Research15, 8125-8148) was introduced immediately 5`of the translation initiation codon and the resulting cDNA was subcloned into the pCAGGS vector (Niwa, H., Yamamura, K. and Miyazaki, J (1991), Gene 8, 193-199) downstream of the chicken β-actin promotor. The IRES derived from encephalmyocarditis virus (Kim, D.G., Kang, H.M., Jang, S.K. and Shin H.S. (1992) Mol.Cell.Biol. 12, 3636-3643) followed by the GFP cDNA containing a Ser65Thr mutation (Heim, R., Cubitt, A.B., Tsien, R.Y. (1995) Nature 373, 663-664) was then cloned 3` to the TRP8b cDNA. The IRES sequence allows the simultaneous translation of TRP8b and GFP from one transcript. Thus, transfected cells can be detected unequivocally by the development of green fluorescence.

For monitoring of the intracellular Ca^{2+} concentration human embryonic kidney (HEK 293) cells were cotransfected with the pcDNA3-TRP8b vector and the pcDNA3-GFPvector in a molar ratio of 4 : 1 in the presence of lipofectamine (Quiagen, Hilden, Germany). To obtain pcDNA3-TRP8b the entire protein coding region of TRP8b including the consensus sequence for initiation of translation in vertebrates (Kozak, M. (1987) Nucleic Acids Research15, 8125-8148) was subcloned into the pcDNA3 vector (Invitrogen, Groningen, Netherlands). Calcium monitoring and patch clamp experiments were carried out two days and one day after transfection, respectively.

(C) Chromosomal localization of the Trp8 gene

The chromosomal localization of the human TRP8 gene was performed using NIGMS somatic hybrid mapping panel No.2 (Coriell Institute, Camden, NJ, USA) previously described (Drwinga, H.L., Toji, L.H., Kim, C.H., Greene, A.E., Mulivor, R.A. (1993) Genomics 16, 311-314; Dubois, B.L. and Naylor, S.L. (1993) Genomics 16, 315-319).

(D) In Vitro Translation, glutathione - sepharose and calmodulin agarose binding assay

N– and C-terminal Trp8-fragments were subcloned into the pGEX-4T2 vector (Amersham Pharmacia Europe, Freiburg, Germany) resulting in glutathione-S-transferase (GST)-Trp8 fusion constructs (Fig. 4). The GST-TRP8-fusion proteins were expressed in E. coli BL 21 cells and purified using glutathione - sepharose beads (Amersham Pharmacia Biotech Europe, Freiburg, Germany).

In vitro translation of human Trp8 cDNA and Xenopus laevis calmodulin cDNA (Davis, T.N. and Thorner, J. Proc.Natl.Acad.Sci. USA 86, 7909-7913.) was performed in the presence of ^{35}S-methionine using the TNT coupled transcription/translation kit (Promega, Madison, USA). Translation products were purified by gel fliltration (Sephadex G50, Amersham Pharmacia Biotech Europe, Freiburg, Germany) and equal amounts of ^{35}S labeled probes were incubated for 2 h with glutathione beads bound to GST - Trp8 or calmodulin - agarose (Calbiochem) in 50 mM Tris-HCl, pH 7.4, 0.1 % Triton X-100, 150 mM NaCl in the presence of 1 mM Ca^{2+} or 2 mM EGTA. After three washes, bound proteins were eluted with SDS sample buffer, fractionated by SDS-PAGE and ^{35}S labeled proteins were detected using a Phosphor Imager (Fujifilm, Tokyo, Japan).

(E) Calcium measurements

The intracellular Ca^{2+} concentration ($[Ca^{2+}]_i$) was determined by dual wavelength fura-2 fluorescence ratio measurements (Tsien, R.Y. (1988) Trends Neurosci. 11, 419-424) using a digital imaging system (T.I.L.L. Photonics, Planegg, Germany). HEK cells were grown in minimal essential medium in the presence of 10 % fetal calf serum and cotransfected with the pcDNA3-TRP8b vector and the pCDNA3-GFPvector as described above (B). Transfected cells were detected by development of green fluorescence. The cells were loaded with 4μM fura-2/AM (Molecular Probes, Oregon, USA) for one hour. After loading the cells were rinsed 3 times with buffer B1 (10 mM Hepes, 115 mM NaCl, 2 mM $MgCl_2$, 5mM KCl, pH 7.4) and the $[Ca^{2+}]_i$ was calculated from the fluorescence ratios obtained at 340 and 380 nm excitation wavelengths as described (Garcia, D.E., Cavalié, A. and Lux, H.D. (1994) J. Neurosci 14, 545-553).

(F) Electrophysiological recordings

HEK cells were transfected with the eucaryotic expression plasmid pdiTRP8 described in (B) and electrophysiolocigal recordings were carried out one day after transfection. Single cells were voltage clamped in the whole cell mode of the patch clamp technique as described (Hamill, O.P., Marty, A., Neher, E., Sakmann, B. and Sigworth, F.J. (1981) Pflügers Arch. 391, 85-100; Philipp, S., Cavalié, A., Freichel, M., Wissenbach, U., Zimmer, S., Trost, C., Marquart, A., Murakami, M. and Flockerzi, V. (1996) EMBO J. 6166-6171). The pipette solution contained contained (mM): 140 aspartic acid, 10 EGTA, 10 NaCl, 1 MgCl2, 10 Hepes (pH 7.2 with CsOH) or 125 CsCl, 10 EGTA, 4 $CaCl_2$ 10 Hepes (pH 7,2 with CsOH). The bath solution contained (mM): 100 NaCl, 10 CsCl, 2 $MgCl_2$, 50 mannitol, 10 glucose, 20

Hepes (pH 7,4 with CsOH) and 2 $CaCl_2$, or no added $CaCl_2$ (-Ca^{2+} solution). Divalent free bath solution contained (mM): 110 N-methyl-D-glucamine (NMDG). Whole cell currents were recorded during 100 msec voltage ramps from –100 to +100 mV at varying holding potentials.

(G) In Situ Hybridization

In situ hybridizations were carried out using formalin fixed tissue slices of 6 - 8 µM thickness. The slices were hydrated and incubated in the presence of PBS buffer including 10 µg / ml proteinase K (Roche Diagnostics, Mannheim, Germany) for 0.5 h. The slices were hybridized at 37°C using biotinylated deoxy-oligonucleotides (0.5 pmol / µl) in the presence of 33 % formamide for 12 h. Furthermore the slices were several times rinsed with 2 x SSC and incubated at 25°C for 0.5 h with avidin / biotinylated horse raddish peroxidase complex (ABC, DAKO, Santa Barbara, USA). After several washes with PBS buffer the slices were incubated in the presence of biotinylated tyramid and peroxide (0.15 % w/v) for 10 min, rinsed with PBS buffer and additionally incubated with ABC complex for 0.5 h. The slices were washed with PBS buffer and incubated in the presence of DAB solution (diaminobenzidine (50µg / ml), 50 mM Tris/EDTA buffer pH 8.4, 0.15 % H_2O_2 in N,N - dimethyl-formamide; Merck, Darmstadt, Germany), The detection was stopped after 4 minutes by incubating the slides in water. Tyramid was biotinylated by incubating NHS-LC Biotin (sulfosuccinimidyl-6-(biotinimid)-hexanoat), 2.5 mg / ml; Pierce, Rockford, USA) and tyramin-HCl (0.75 mg / ml, Sigma) in 25 mM borate buffer pH 8.5 for 12 h. The tyramid solution was diluted 1 - 5 : 1000 in PBS buffer.

(H) GenBank accession numbers: TRP8a, Aj243500; TRP8b Aj243501

Example 2: Expression of TRP8 transcripts

In search of proteins distantly related to the TRP family of ion channels, an human expressed sequence tag (EST, GenBank accession number 1404042) was identified in the GenBank database using BLAST programms (at the National Center for Biotechnology Information (NCBI); Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J.J. (1990) Mol. Biol. 5, 403-410) being slightly homologous to the VR1 gene. Several human placenta cDNA libraries were constructed and screeened with this EST DNA as probe. Several full length

cDNA clones were identified and isolated. The full length cDNA clones encoded two putative proteins differing in three amino acids and were termed Trp8a and Trp8b (Fig. 1c, 2a, 7 and 8A). This finding was reproduced by isolating cDNA clones from two cDNA libraries constructed from two individual placentas. The derived protein sequence(s) comprises six transmembrane domains, a characteristic overall feature of trp channels and related proteins (Fig.: 1b). The sequence is closely related to the meanwhile published calcium uptake transport protein 1 (CaT1), isolated from rat intestine (Peng, J.B., Chen, X.Z., Berger, U.V., Vassilev, P.M., Tsukaguchi, H., Brown, E.M. and Hediger M.A.(1999) J Biol Chem. 6;274, 22739-22746) and to the epithelial calcium uptake channel (ECaC) isolated from rabbit kidney (Hoenderop, J.G., van der Kemp, A.W., Hartog, A., van de Graaf, S.F., van Os, C.H., Willems, P.H. and Bindels, R.J. (1999) J Biol Chem. 26;274, 8375-8378). Expression of Trp8a/b transcripts are detectable in human placenta, pancreas and prostate (Fig.: 5) and the size of the Northern signal (3.0 kb) corresponds with the size of the isolated full length cDNAs. In addition, a shorter transcript of 1.8 kb, probably a splice variant, is detectable in human testis. The Trp8 mRNA is not expressed in small intestine or colon (Fig.: 5) implicating that Trp8 is not the human ortholog of the rat CaT1 or rabbit ECaC proteins. To investigate whether there are other related sequences Trp8a/b derived primers (UW241, 5'-TAT GAG GGT TCA GAC TGC-3' and UW242, 5'-CAA AGT AGA TGA GGT TGC-3') were used to amplify a 105 bp fragment from human genomic DNA being 95% identical on the nucleotide level to the Trp8 sequence (data not shown). This indicates the existence of several similar sequences in humans at least at the genomic level.

Example 3: Two variants of the Trp8 protein (Trp8a and Trp8b) arise by polymorphism

Two variants of the Trp8 cDNA were isolated from human placenta (Fig.: 2A, 7 and 8A) which encoded two proteins which differ in three amino acids and were termed Trp8a and Trp8b. Trp8a/b specific primers were designed to amplify a DNA fragment of 458 bp of the Trp8 gene from genomic DNA isolated from human T-lymphocytes (primer pair: UW243, 5'-CAC CAT GTG CTG CAT CTA CC-3' and UW244, 5'-CAA TGA CAG TCA CCA GCT CC-3'). The amplification product contains a part of the sequence where the derived protein sequence of Trp8a comprises the amino acid valine and the Trp8b sequence methionine as well as a silent base pair exchange (g versus a) and an intron of 303bp (Fig.: 2.A, B). Both variants of the Trp8 genes (a,b) were amplified from genomic DNA in equal amounts indicating the existence of both variants in the human genome and therefore being not the

result of RNA editing (Fig.: 2B). The Trp8a gene can be distinguished from the Trp8b gene by cutting the genomic fragment of 458bp with the restriction enzyme Bsp1286I (Fig. 2B). Using human genomic DNA isolated from blood of twelve human subjects as template, the 458bp fragment was amplified and restricted with BSP1286I. In eleven of the tested subjects only the Trp8b gene is detectable, while one subject (7) contains Trp8a and Trp8b genes (Fig.: 2D). These implicates that the two Trp8 variants arise by polymorphism and do not represent individual genes. Using Trp8 specific primers and chromosomal DNA as template, the Trp8 locus is detectable on chromosome 7 (Fig.: 2C).

Example 4: Trp8b is a calcium permeable channel

The protein coding sequence of the Trp8b cDNA was subcloned into pcDNA3 vector (Invitrogen, Groningen, Netherlands) under the control of the cytomegalovirus promotor (CMV). Human embryonic kidney (HEK 293) cells were cotransfected with the Trp8b pcDNA3 construct (pcDNA3-Trp8b vector) and the pcDNA3-GFPvector encoding the green fluorescent protein (GFP) in 4:1 ratio. The Trp8b cDNA and the cDNA of the reporter, GFP, was transiently expressed in human embryonic kidney (HEK 293) cells. The intracellular Ca^{2+} concentration ($[Ca^{2+}]_i$) and changes of $[Ca^{2+}]_i$ were determined by dual wavelength fura-2 fluorescence ratio measurements (Fig.: 3F) in cotransfected cells which were identified by the green fluorescence of the reporter gene GFP.

Dual wavelength fura-2 fluorescence ratio measurement is a standard procedure (e.g. in: An introduction of Molecular Neurobiology (ed. Hall, Z.W.)Sinauer Associates, Sunderland, USA (1992)) using fura-2, which is a fluorescent Ca^{2+} sensitive dye and which was designed by R.Y.Tsien (e.g. Trends Neurosci. 11, 419-424 (1988) based upon the structure of EGTA. Its fluorescence emission spectrum is altered by binding to Ca^{2+} in the physiological concentration range. In the absence of Ca^{2+}, fura-2 fluoresces most strongly at an excitation wavelength of 385 nm; when it binds Ca^{2+}, the most effective excitation wavelength shifts to 345 nm. This property is used to measure local Ca^{2+} concentrations within cells. Cells can be loaded with fura-2 esters (e.g. fura-2AM) that diffuse across cell membranes and are hydrolyzed to active fura-2 by cytosolic esterases.

In the presence of 1mM Ca^{2+}, Trp8 expressing cells typically contained more than 300 nM cytosolic Ca^{2+}, while non transfected controls contained less than 100 nM Ca^{2+} ions (Fig. 3F).

When Trp8b transfected cells were incubated without extracellular Ca^{2+}, the intracellular Ca^{2+} concentration ($[Ca^{2+}]_i$) decreased to levels comparable to non transfected cells. Readdition of 1mM Ca^{2+} to the bath resulted in significant increase of the cytosolic $[Ca^{2+}]$ in Trp8b transfected cells, but not in controls (Fig.: 3F). After readdition of Ca^{2+} ions to the bath solution, the cytosolic Ca^{2+} concentration remains on a high steady state level in the Trp8b transfected cells.

Example 5: Trp8 expressing cells show calcium selective inward currents

To characterize in detail the electrophysiological properties of TRP8, TRP8 and GFP were coexpressed in HEK293 cells using the dicistronic expression vector pdiTRP8 and measured currents using the patch clamp technique in the whole cell mode (Hamill, O.P., Marty, A., Neher, E., Sakmann, B. and Sigworth, F.J. (1981) Pflugers Arch., 391, 85-100).

The eucaryotic expression plasmid pdiTRP8 contains the cDNA of Trp8b under the control of the chicken ß-actin promotor followed by an internal ribosome entry side (IRES) and the cDNA of the green fluorescent protein (GFP). To obtain pdiTRP8 carrying the entire protein coding regions of TRP8b and the GFP (Prasher, D.C. et al. (1992), Gene 111, 229-233), the 5`and 3´-untranslated sequences of the TRP8b cDNA were removed, the consensus sequence for initiation of translation in vertebrates (Kozak, M. (1987) Nucleic Acids Research15, 8125-8148) was introduced immediately 5`of the translation initiation codon and the resulting cDNA was subcloned into the pCAGGS vector (Niwa, H., Yamamura, K. and Miyazaki, J (1991), Gene 8, 193-199) downstream of the chicken β-actin promotor. The IRES derived from encephalmyocarditis virus (Kim, D.G., Kang, H.M., Jang, S.K. and Shin H.S. (1992) Mol.Cell.Biol. 12, 3636-3643) followed by the GFP cDNA containing a Ser65Thr mutation (Heim, R., Cubitt, A.B., Tsien, R.Y. (1995) Nature 373, 663-664) was then cloned 3` to the TRP8b cDNA. The IRES sequence allows the simultaneous translation of TRP8b and GFP from one transcript. Thus, transfected cells can be detected unequivocally by the development of green fluorescence.

In the presence of 2 mM external calcium, Trp8b transfected HEK cells show inwardly rectifying currents, the size of which depends on the level of intracellular calcium and the electrochemical driving force. The resting membrane potential was held either at –40 mV, or, to lower the driving force for calcium influx in between pulses, at + 70 mV. Current traces

were recorded in response to voltage ramps from –100 to +100 mV, that were applied every second. To monitor inward and outward currents over time, we analyzed the current size at – 80 and + 80 mV of the ramps. Figure 3A shows a representative trace of the current at – 80 mV over time. Both at a holding potential of –40 mV or at +70 mV, the currents are significantly larger than in cells transfected with only the GFP containing vector (Fig.: 3E). Interestingly, after changing to a positive holding potential, current size in Trp8 transfected cells slowly increases and reaches steady state after approximately 70 seconds (Fig.: 3A). To determine the selectivity of the induced currents, we then perfused the cells with solutions that either contain no sodium, no added Ca^{2+} (Fig. 3A, C) or a sodium containing, but divalent ion free bath solution. To control for the effect of the solution change alone, we also perfused with normal bath (see puff in Fig. 3A). While removal of external Ca^{2+} completely abolishes the trp 8 induced currents - the remaining current being identical in size and shape to the control (Fig.: 3A, C, E), removal of external sodium has no effect (Fig.: 3E). An important hallmark of calcium selective channels (e.g. Vennekens, R., Hoenderop, G.J., Prenen, J., Stuiover, M., Willems, PHGM, Droogmans, G., Nilius, B.and Bindels, R.J.M (1999) J. Biol. Chem. 275, 3963-3969), is their ability to conduct sodium only if all external divalent ions, namely Ca^{2+} and magnesium are removed. To test whether the trp 8 channel conforms with this phenomenon normal bath solution was switched to a solution containing only sodium and 1 mM EGTA. As can be seen in Figure 3B and D, Trp8 transfected cells can now conduct very large sodium currents. Interestingly, immediately after the solution change, the currents first become smaller before increasing rapidly, indicating that the pore may initially still be blocked by calcium a phenomenon usually called anomalous mole fraction behaviour (Warnat, J., Philipp, S., Zimmer, S.,.Flockerzi, V., and Cavalié A.(1999) J.Physiol. (Lond) 518, 631-638). The measured outward currents of Trp8 transfected cells in normal bath solution are not significantly different from non-transfected control cells or cells which only express the reporter gene GFP. As the removal of external Ca^{2+} abolishes the Trp8 specific current, the remaining current was subtracted from the current before the solution change to obtain the uncontaminated Trp8 conductance (see inset in Fig.: 3C). As expected from the given ionic conditions (high EGTA inside, 2 mM Ca^{2+} outside), the current-voltage relationship now shows prominent inward rectification with little to no outward current.

Both the time course of the development of Trp8 currents and the size of the currents depend on the frequency of stimulation (data not shown), the internal and external Ca^{2+} concentration

and the resting membrane potential, suggesting that Trp8 calcium conductance is intrically regulated by a Ca^{2+} mediated feedback mechanisms.

Example 6: Ca^{2+} / calmodulin binds to the C-terminus of the Trp8 protein

To test whether calmodulin, a prime mediator of calcium regulated feedback, is involved, first it was investigated biochemically whether Trp8 protein can bind calmodulin. Trp8 cDNA was in vitro translated in the presence of ^{35}S-methionine and the product incubated with calmodulin-agarose beads. After several washes either in the presence or abscence of Ca^{2+}, the beads were incubated in Laemmli buffer and subjected to SDS-polyacrylamide gel electrophoresis. In the presence of Ca^{2+} (1mM), but not in the absence of Ca^{2+}, Trp8 protein binds to calmodulin (Fig.: 4B).

To narrow down the binding site, two approaches were undertaken: Firstly, GST-TRP8 fusion proteins of various intracellular domains of Trp8 were constructed, expressed in E. coli and bound to gluthathione sepharose beads. These beads were then incubated with in vitro translated ^{35}S- labeled calmodulin, washed and subjected to gel electrophoresis. Secondly, truncated versions of in vitro translated Trp8 protein were used in the above described binding to calmodulin-agarose. As shown in Figure 4A, and C, fusion proteins of the N-terminal region (N1, N2) of Trp8 did not bind calmodulin, while C-terminal fragments (C1, C2, C3, C4) showed calmodulin binding in the presence of calcium (for localization of fragments within the entire Trp8 protein see Fig. 4C). Accordingly, a truncated version of in vitro translated Trp8, which lacks the C-terminal 32 amino acid residues did not bind to calmodulin-agarose (4B). We have restricted the calmodulin binding site to amino acid residues 691 to 711 of the Trp8 protein. This calmodulin binding site does not resemble the typical conserved IQ - motif of conventional myosins, but has limited sequence homology to the calcium dependent calmodulin binding site 1 of the transient receptor potential like (trpl) protein of Drosophila melanogaster (Warr and Kelly, 1996) with several charged amino acid residues conserved. The sequence of the calmodulin binding site of the Trp8 protein resembles a putative amphipathic α-helical wheel structure with a charged and a hydrophobic site according to a model proposed by Erickson-Vitanen and De Grado (1987, Methods Enzymol. 139, 455-478.).

Example 7: Expression of Trp8 transcripts in human placenta and pancreas

Several slides from a human placenta of a ten week old abort were used for in situ hybridization experiments. The in situ hybridization experiments revealed expression of Trp8 transcripts in human placenta (Fig.: 5B). Expression was detectable in trophoblasts and syncytiotrophoblasts of the placenta, but not in Langhans cells.

Trp8 transcripts are detectable in human pancreas (Fig.: 5A). Therefore Trp8 probes were hybridized to tissue sections of human pancreas. The pancreatic tissues were removed from patients with pancreas cancer. Trp8 expression is detectable in pancreatic acinar cells, but not in Langerhans islets (Fig.: 5C). No Trp8 expression was found in regions of pancreatic carcinomas (data not shown).

Furthermore, the Trp8 cDNA is not detectable in human colon nor in human kidney by in situ hybridization as well as by Northern analysis (Fig.: 5A, D). The Northern results taken together with the in situ expression data indicate that the Trp8 protein is not the human ortholog of the CaT1 and ECaC channels cloned from rat intestine (Peng, J.B., Chen, X.Z., Berger, U.V., Vassilev, P.M., Tsukaguchi, H., Brown, E.M. and Hediger M.A.(1999) J Biol Chem. 6;274, 22739-22746) and from rabbit kidney (Hoenderop, J.G., van der Kemp, A.W., Hartog, A., van de Graaf, S.F., van Os, C.H., Willems, P.H. and Bindels, R.J. (1999) J Biol Chem. 26;274, 8375-8378), respectively. Trp8 is unlikely to represent the human version of CaT1 as its expression is undetectable in the small intestine and colon tissues where CaT1 is abundantly expressed. If, however, Trp8 is the human version of rat CaT1, a second gene product appears to be required for Ca^{2+} uptake in human small intestine and colon attributed to CaT1 in rat small intestine and colon.

Example 8: Differential expression of Trp8 transcripts in benign and malign tissue of the prostate

The Trp8 transcripts are expressed in human prostate as shown by hybridization of a Trp8 probe to a commercial Northern blot (Clontech, Palo Alto, USA) (Fig.: 5A). Trp8 transcripts were not detectable by Northern blot analysis using pooled mRNA of patients with benign prostatic hyperplasia (BPH) (Fig.: 5A, prostate*). To examine Trp8 expression on the cellular

level, sections of prostate tissues were hybridized using Trp8 specific cDNA probes (Table 3). Expression of Trp8 transcripts is not detectable in normal prostate (n = 3), benign hyperplasia (BPH, n = 15) or prostatic intraepithelial neoplasia (PIN, n = 9) (Fig.: 6A, C, E). Trp8 transcripts were only detectable in prostate carcinoma (PCA), although with different expression levels. Low expression levels were found in primary carcinomas (2 - 10 % of the carcinoma cells, n = 8) (Fig.: 7B) . Much stronger expression was detectable in rezidive carcinoma (10 - 60 %) (Fig.: 7D, n = 6) and metastases of the prostate (60 - 90 %, n = 4) (Fig.: 7F). Thus it has to be concluded that the commercial Northern blot used in Fig.: 5A contains not only normal prostate mRNA as indicated by the distributor. According to the distributors instructions the prostate mRNA used for this Northern blot was collected from 15 human subjects in the range of 14 to 60 years of age. This prostate tissue was not examined by pathologic means. Since Trp8 expression is not detectable in normal or benign prostate, this finding implicates that the mRNA used for this Northern blot was extracted in part from prostatic carcinoma tissue. To summarize, Trp8 expression is only detectable in malign prostate and, thus, the Trp8 cDNA is a marker for prostate carcinoma. The results are summarized in Table 4.

Table 3

Trp8 probes used for in situ hybridization:

Probes (antisense)

1.) 5' TCCGCTGCCGGTTGAGATCTTGCC 3'

2.) 5' CTTGCTCCATAGGCAGAGAATTAG 3'

3.) 5' ATCCTCAGAGCCCCGGGTGTGGAA3'

Controls (sense)

1.) 5' GGCAAGATCTCAACCGGCAGCGGA 3'

2.) 5' CTAATTCTCTGCCTATGGAGCAAG 3'

3.) 5' TTCCACACCCGGGGCTCTGAGGAT 3'

Table 4

Prostate	total	negative	positive
normal	3	3	0
BPH	15	15	0
PIN	9	9	0

carcinoma	18	1	17

(B) Differential expression of Trp8 transcripts in benign and malign tissue of the uterus

Moreover it could be shown that Trp8 is expresed in endometrial cancer (also called cancer of the uterus, to be distinguished from uterine sarcoma or cancer of the cervix) whereas no expression was observed in normal uterus tissue. Thus, Trp8 also is a specific marker for the diagnosis of the above cancer (Fig. 12).

Example 9: Characterization of Trp9

The complete protein coding sequence of Trp9 was determined (Fig. 9). Trp 9 transcripts are predominantly expressed in the human prostate and in human colon. As it could be shown by Northern blot analysis, there is no difference of the expression of TRP9 in benigne prostata hyperplasia (BPH, Fig. 13, upper panel left) or prostate carcinoma (Fig. 13, upper panel right). However, Trp9 is useful as a reference marker for prostata carcinoma, i.e. can be used for quantifying the expression level of Trp8. The ratio of the expression of Trp8:Trp9 in patients and healthy individuals is useful for the development of a quantitative assay.

Example 10: Characterization of Trp10

The complete protein coding sequence of TRP10 (a and b) was determined by biocomputing (Fig. 10 and 11). Using a 235 bp fragment of the Trp10 cDNA as probe in Northern blot analysis TRP10 transcripts could only be detected in mRNA isolated from individuals with prostate cancer (Fig. 13, bottom panel) but not in mRNA isolated from benign tissue of the prostate (prostate BPH) nor in mRNA isolated from heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas. The 235 bp cDNA fragment of the Trp10 cDNA was amplified using the primer pair UW248 5'-ACA GCT GCT GGT CTA TTC C-3' and UW249 5'-TAT

GTG CCT TGG TTT GTA CC-3' and prostate cDNA as template. In summary, Trp10a and Trp10b , like TRP8 are also expressed in malignant prostate tissue. So far, its expression could not be observed in any other tissue examined (see above). Thus, Trp 10a and Trp10b are also useful markers which are specific for malignant prostate tissue.

Furthermore, database searches in public databases of the national center for biological information (NCBI) revealed the existence of several expressed sequence tags (EST clones) being in part identical to the Trp10 sequence. These EST clones were originally isolated from cancer tissues of lung, placenta, prostate and from melanoma. These clones include the clones with the following accession numbers: BE274448, BE408880, BE207083, BE791173, AI671853, BE390627. The results demonstrate that cancer cells of these tissues express Trp10 related transcripts whereas no expression of Trp10 transcripts in the corresponding healthy tissues are detectable (Figure 13). Furthermore, it could be shown that in cancer cells of melanoma and prostate cancer Trp10 transcripts are expressed as shown by in situ hybridizations using 4 antisense probes (Figure 14A – E and 13K-O and Table 2, above). Furthermore, it could clearly be shown that cancer cells of these tissues expressing Trp10 transcripts also express Trp10-antisense transcripts as shown in Figure 14F-J, Figure 14P-R and Figure14T by in situ hybridizations using 4 sense probes (Table 2, above). The in situ hybridization experiments demonstrate that detection of a subset of cancer cells derived from carcinoma of lung, placenta, prostate and melanoma is feasible using antisense as well as sense probes complementary to Trp10 transcripts or complementary to Trp10-antisense transcripts, respectively.

The foregoing is meant to illustrate but not to limit the scope of the invention. The person skilled in the art can readily envision and produce further embodiment, based on the above teachings, without undue experimentation.

What Is claimed Is:

1. An isolated nucleic acid molecule encoding the human prostate carcinoma associated protein Trp8a, Trp8b, Trp9, Trp10a or Trp10b or a protein exhibiting biological properties of Trp8a, Trp8b, Trp9, Trp10a or Trp10b and being selected from the group consisting of

(a) a nucleic acid molecule encoding a protein that comprises the amino acid sequence depicted in Figure 7, 8A, 9, 10 or 11;

(b) a nucleic acid molecule comprising the nucleotide sequence depicted in Figure 7, 8A, 9, 10 or 11;

(c) a nucleic acid molecule included in DSMZ Deposit No. DSM 13579, DSM 13580, DSM 13584 , DSM 13581 or DSM....;

(d) a nucleic acid molecule which hybridizes to a nucleic acid molecule specified in (a) to (c);

(e) a nucleic acid molecule the nucleic acid sequence of which deviates from the nucleic sequences specified in (a) to (d) due to the degeneration of the genetic code; and

(f) a nucleic acid molecule, which represents a fragment, derivative or allelic variation of a nucleic acid sequence specified in (a) to (e).

2. A recombinant vector containing the nucleic acid molecule of claim 1

3. The recombinant vector of claim 2 wherein the nucleic acid molecule is operatively linked to regulatory elements allowing transcription and synthesis of a translatable RNA in prokaryotic and/or eukaryotic host cells.

4. A recombinant host cell which contains the recombinant vector of claim 3.

5. The recombinant host cell of claim 4, which is a mammalian cell, a bacterial cell, an insect cell or a yeast cell.

6. An isolated protein exhibiting biological properties of the human prostate carcinoma associated protein Trp8a, Trp8b, Trp9, Trp10a or Trp10b which is encoded by a nucleic acid molecule of claim 1.

7. A recombinant host cell that expresses the isolated protein of claim 6.

8. A method of making an isolated protein exhibiting biological properties of the human prostate carcinoma associated protein Trp8a, Trp8b, Trp9, Trp10a or Trp10b comprising:
(a) culturing the recombinant host cell of claim 6 under conditions such that said protein is expressed; and
(b) recovering said protein.

9. The protein produced by the method of claim 8.

10. An antisense RNA sequence characterized in that it is complementary to an mRNA transcribed from a nucleic acid molecule of claim 1 or a part thereof and can selectively bind to said mRNA or part thereof, said sequence being capable of inhibiting the synthesis of the protein encoded by said nucleic acid molecule.

11. A ribozyme characterized in that it is complementary to an mRNA transcribed from a nucleic acid molecule of claim 1 or a part thereof and can selectively bind to and cleave said mRNA or part thereof, thus inhibiting the synthesis of the protein encoded by said nucleic acid molecule.

12. An inhibitor characterized in that it can suppress the activity of the protein of claim 6.

13. A method for diagnosing a prostate carcinoma which comprises contacting a target sample suspected to contain the protein Trp8a, Trp8b, Trp10a and/or Trp10b or the Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA with a reagent which reacts with Trp8a, Trp8b, Trp10a and/or Trp10b or the Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA and detecting Trp8a, Trp8b, Trp10a and/or Trp10b or the Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA.

14. The method of claim 13, wherein the reagent is a nucleic acid.

15. The method of claim 13, wherein the reagent is an antibody.

16. The method of claim 13, wherein the reagent is detectably labeled.

17. The method of claim 16, wherein the label is selected from the group consisting of a radioisotope, a bioluminescent compound, a chemiluminescent compound, a fluorescent compound, a metal chelate, or an enzyme.

18. A method for diagnosing an endometrial cancer (carcinoma of the uterus) which comprises contacting a target sample suspected to contain the protein Trp8a and/or Trp8b or the Trp8a and/or Trp8b encoding mRNA with a reagent which reacts with Trp8a and/or Trp8b or the Trp8a and/or Trp8a and/or trp8b encoding mRNA and detecting Trp8a and/or Trp8b or the Trp8a and/or Trp8b encoding mRNA.

19. The method of claim 18, wherein the reagent is a nucleic acid.

20. The method of claim 18, wherein the reagent is an antibody.

21. The method of claim 18, wherein the reagent is detectably labeled.

22. The method of claim 21, wherein the label is selected from the group consisting of a radioisotope, a bioluminescent compound, a chemiluminescent compound, a fluorescent compound, a metal chelate, or an enzyme.

23. A method for diagnosing a melanoma, chorion carcinoma, cancer of the lung and of the prostate in a tissue of a subject, comprising contacting a sample with a reagent which detects Trp10a and/or Trp10b antisense RNA or Trp10a and/or Trp10b related antisense RNA.

24. A method for preventing, treating, or ameliorating a prostate tumor, endometrial cancer (carcinoma of the uterus) tumor, a chorion carcinoma, cancer of the lung or melanoma, which comprises administering to a mammalian subject a therapeutically effective amount of a reagent which decreases or inhibits expression of Trp8a, Trp8b, Trp10a and/or Trp10b and/or the activity of Trp8a, Trp8b, Trp10a and/or Trp10b.

25. The method of claim 24, wherein the reagent is a nucleotide sequence comprising an antisense RNA.

26. The method of claim 24, wherein the reagent is a nucleotide sequence comprising a ribozyme.

27. The method of claim 24, wherein the reagent is an inhibitor of Trp8a, Trp8b, Trp10a and/or Trp10b.

28. The method of claim 27, wherein the reagent is an anti-Trp8a-, anti Trp8b-, anti-Trp10a- and/or anti-Trp10b antibody or a fragment thereof.

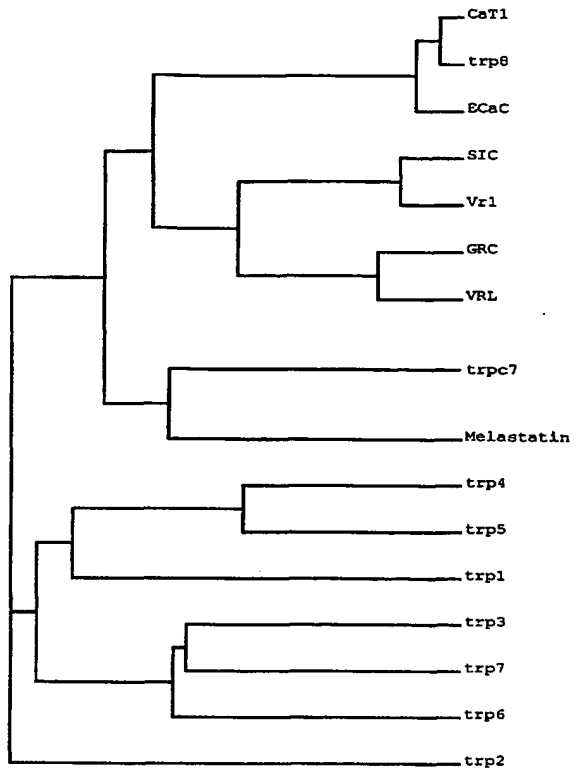
29. A diagnostic kit useful for the detection of Trp8a, Trp8b, Trp10a and/or Trp10b or Trp8a, Trp8b, Trp10a and/or Trp10b encoding mRNA or Trp10a and/or Trp10b antisense transcripts in a sample, wherein the presence of an increased concentration of Trp8a, Trp8b, Trp9, Trp10a and/or Trp10b or Trp8a, Trp8b, Trp9, Trp10a and/or Trp10b encoding mRNA or Trp10a and/or Trp10b antisense transcripts is indicative for a prostate tumor, endometrial cancer (cancer of the uterus) tumor, a chorion carcinoma, cancer of the lung or melanoma, said kit comprising a probe for detection of Trp8a, Trp8b, Trp9, Trp10a or Trp10b or Trp8a, Trp8b, Trp9, Trp10a and/or Trp10b encoding mRNA or Trp10a and/or Trp10b antisense transcripts.

30. The kit of claim 29, wherein the target component to be detected is Trp8a, Trp8b, Trp9, Trp10a and/or Trp10b and the probe is an antibody.

31. A method for identifying a compound which acts as an agonist or antagonist on the ion channels Trp8, Trp9 and/or Trp10, said method comprising contacting a test compound with the ion channel Trp8, Trp9 and/or Trp10,and determining whether said test compound affects the calcium uptake.

Figs. 1A and 1B

A



B

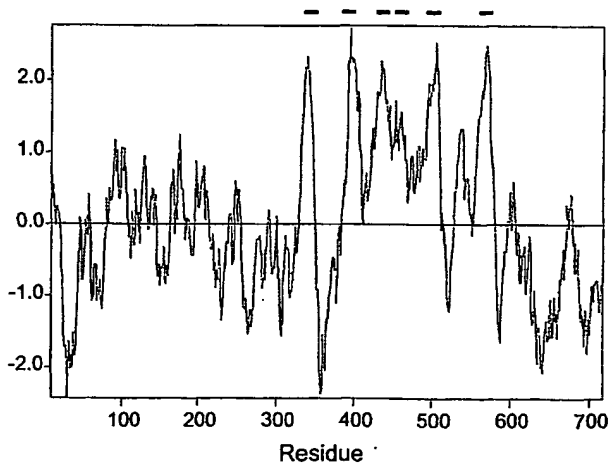
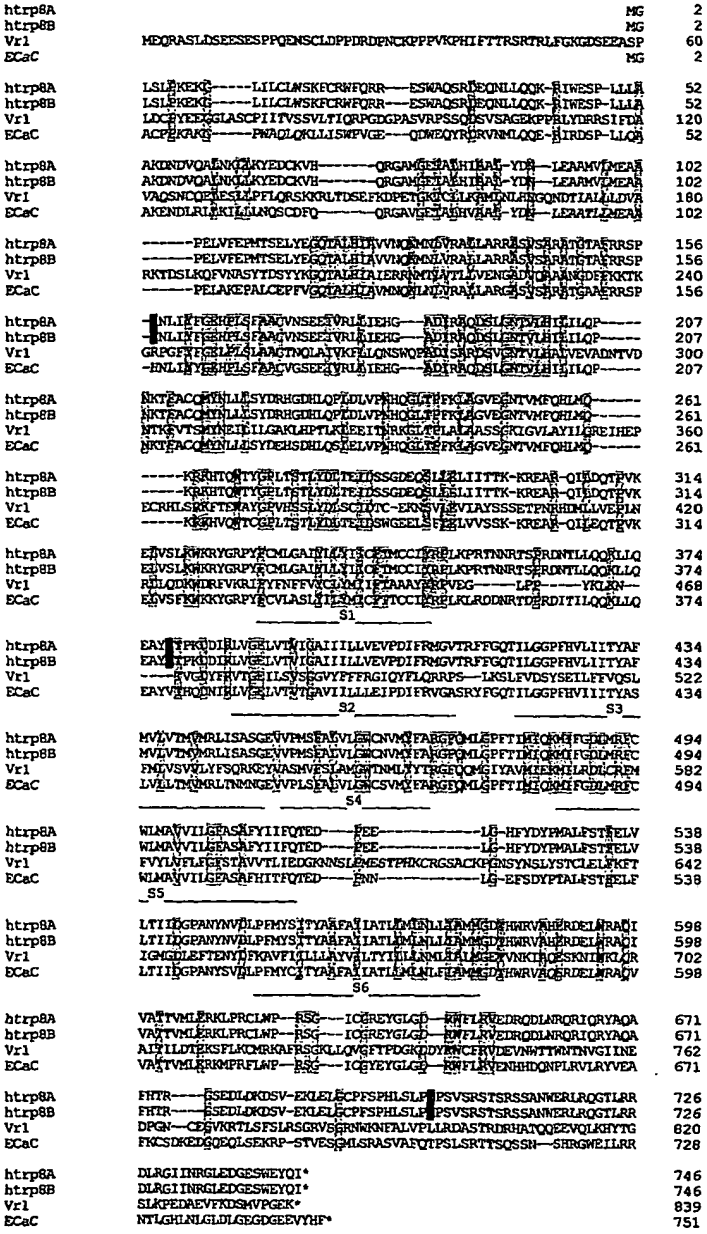
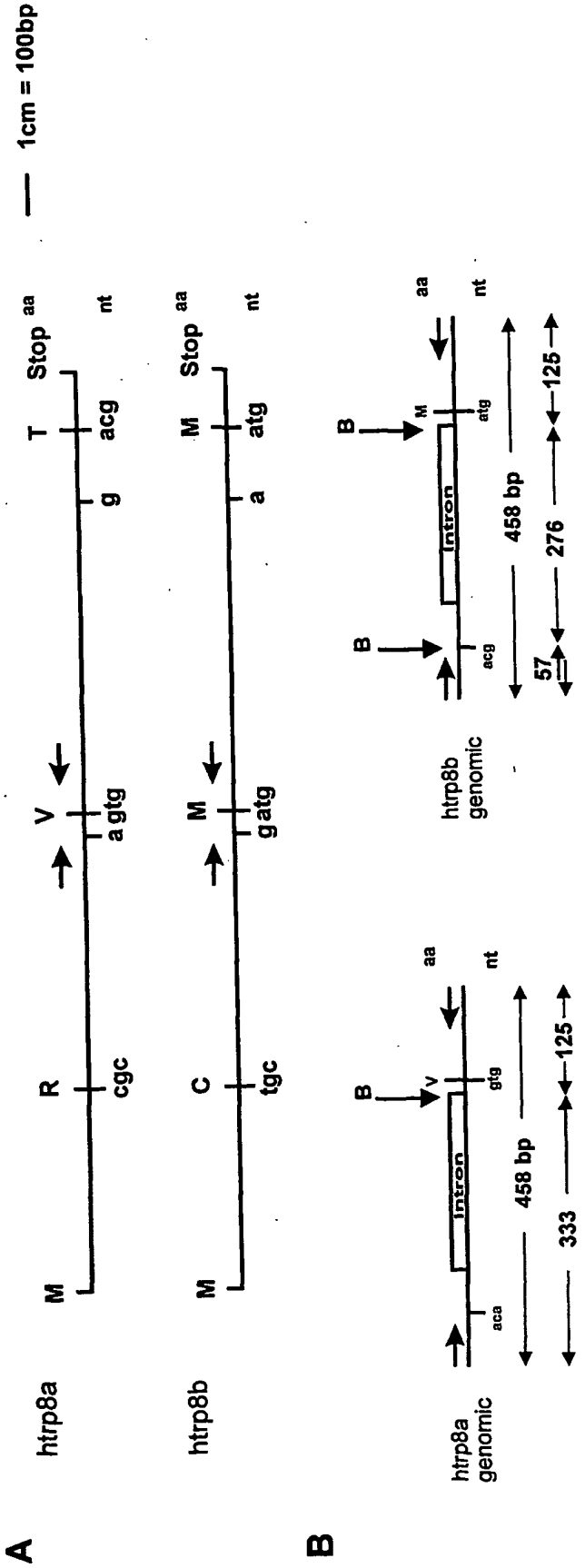


Fig. 1C

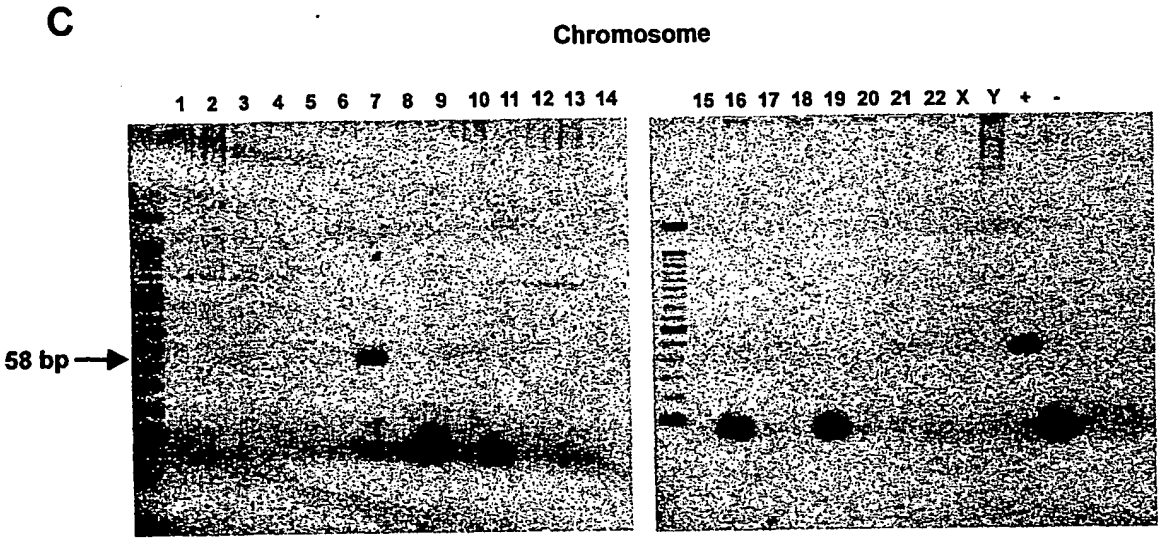
C

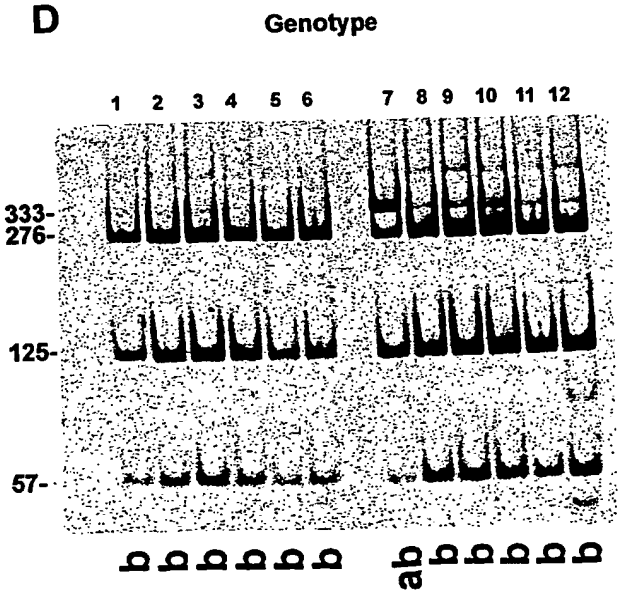




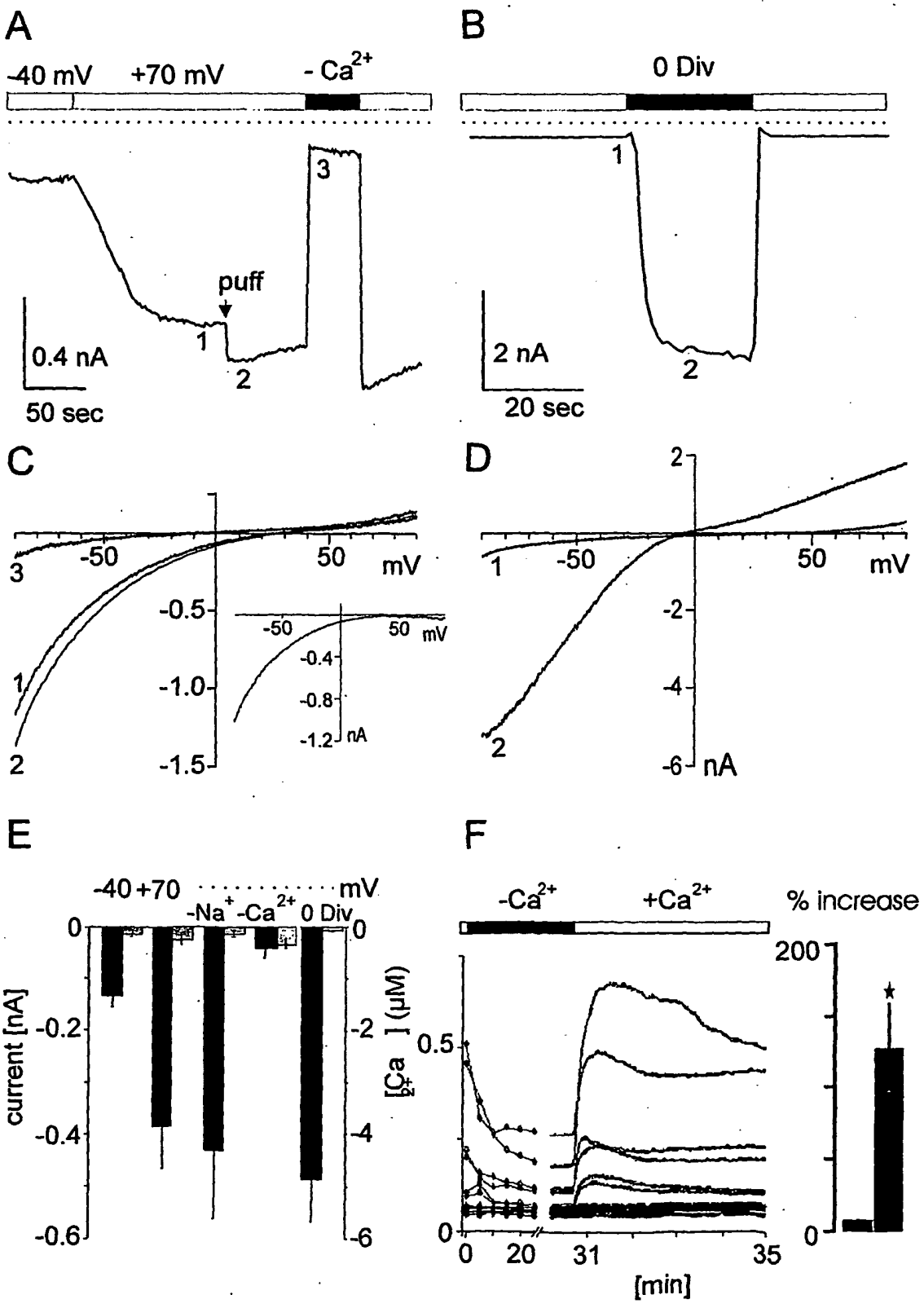
Figs. 2A and 2B

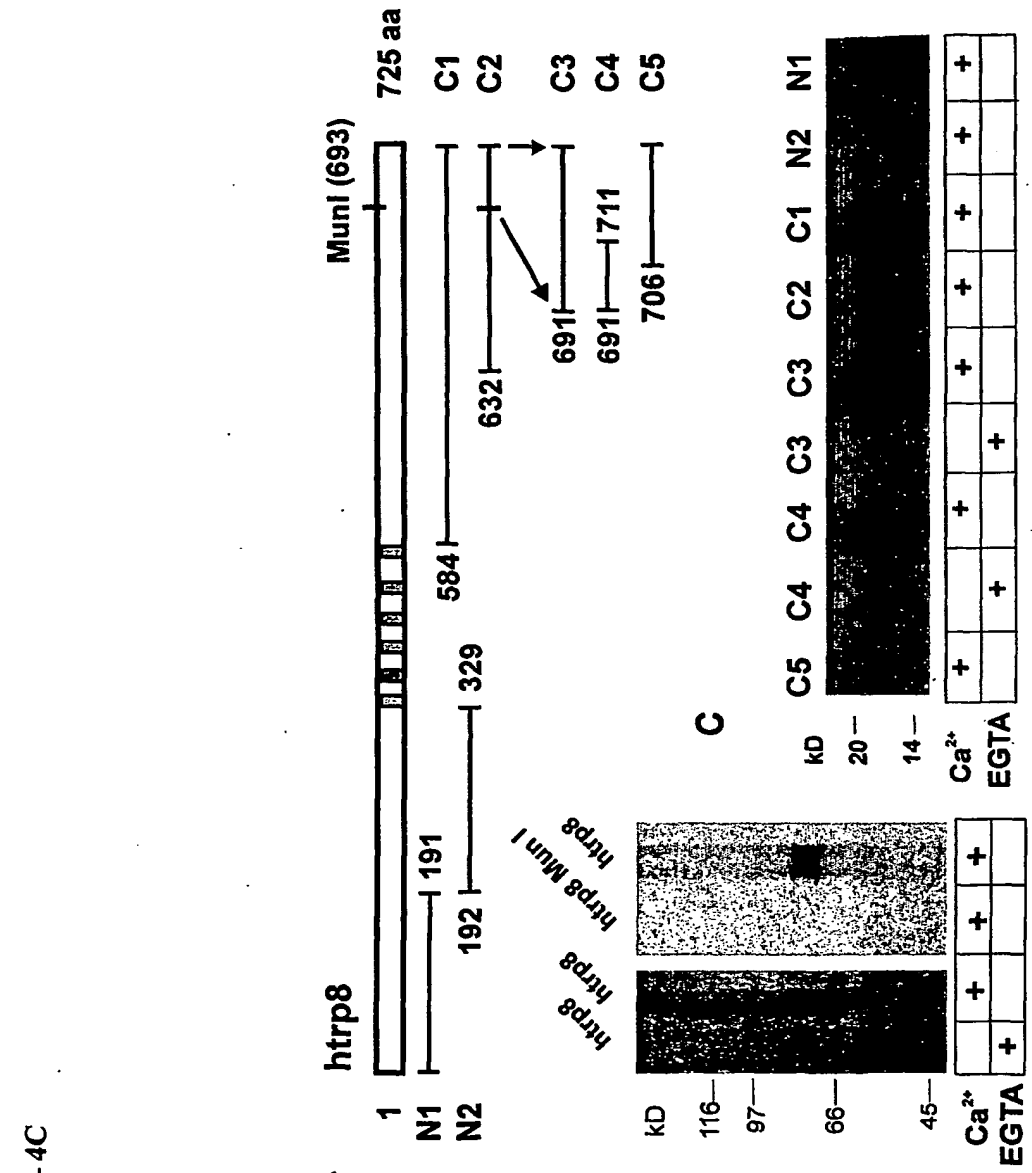
Figs. 2C and 2D





Figs. 3A – 3F


A
-40 mV
+70 mV
- Ca2+
puff
0.4 nA
50 sec
B
0 Div
2 nA
20 sec
C
-50
50
mV
-0.5
-1.0
-1.5
-0.4
-0.8
-1.2
nA
D
-2
-4
-6
nA
E
-40 +70
mV
-Na+ -Ca2+ 0 Div
current [nA]
-0.2
-0.4
-0.6
-2
-4
-6
F
-Ca2+
+Ca2+
% increase
200
0.5
[Ca2+] (µM)
0
20
31
35
[min]


A
htrp8
1
725 aa
MunI (693)
N1
191
584
C1
N2
192
329
632
C2
691
C3
691
711
C4
706
C5
B
htrp8
htrp8 MunI
kD
116
97
66
45
Ca2+
EGTA
C
C5 C4 C4 C3 C3 C2 C1 N2 N1
kD
20
14
Ca2+
EGTA

Figs. 4A – 4C

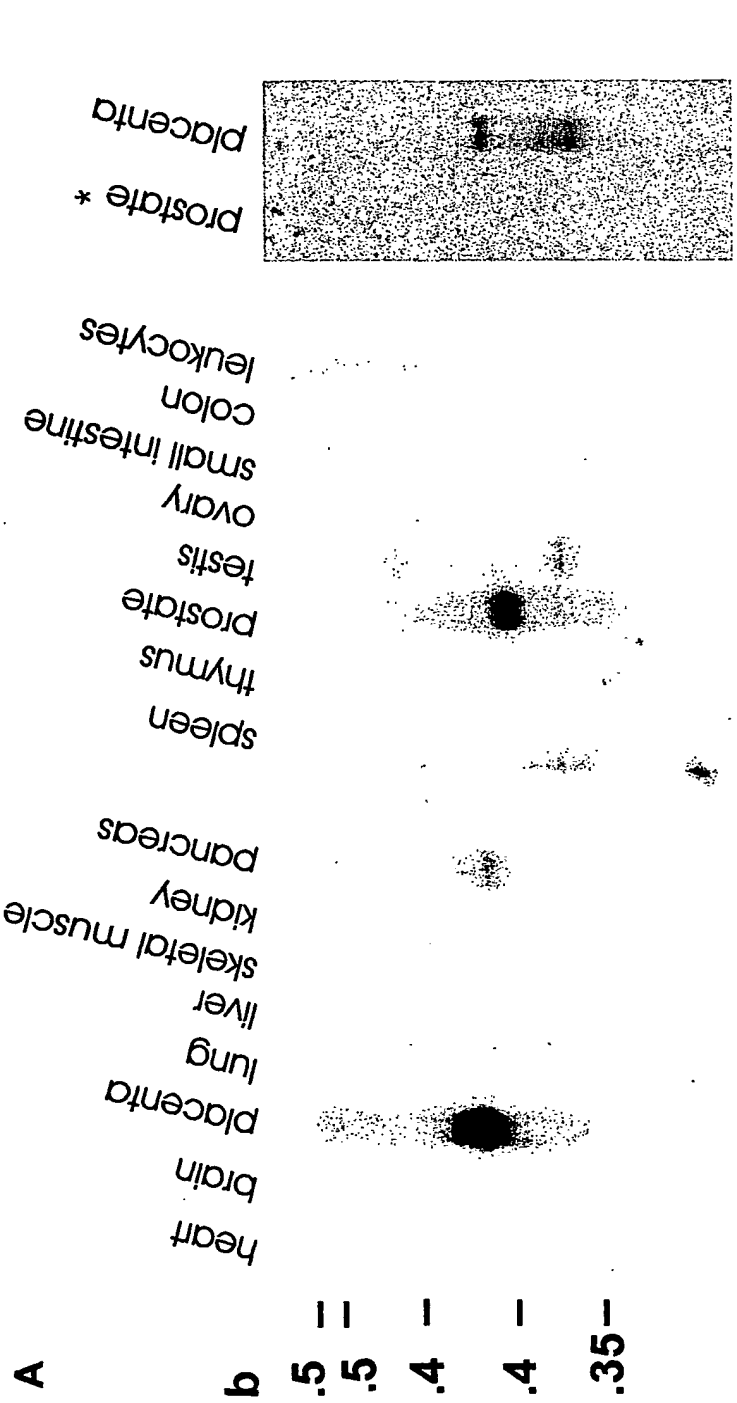
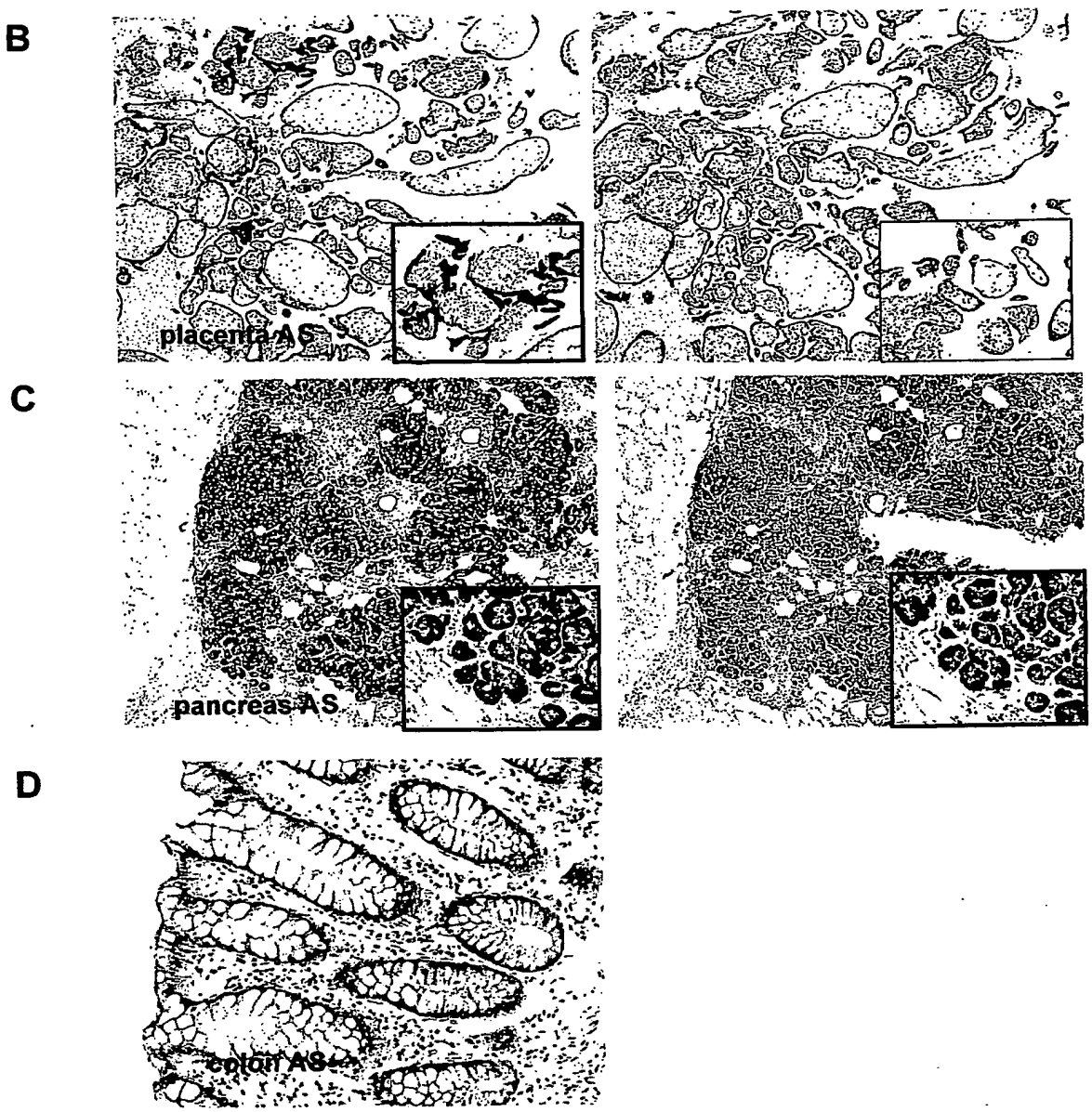
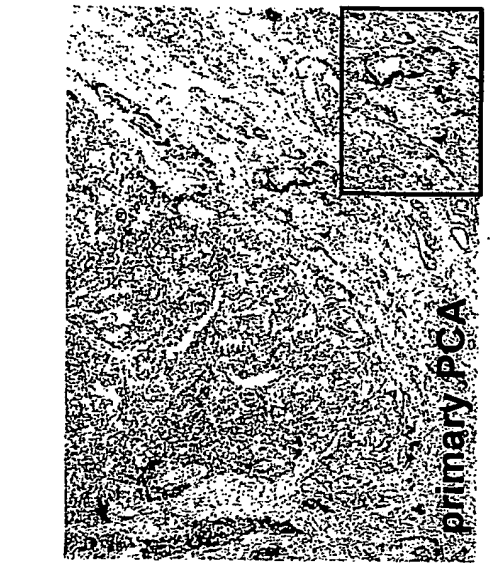
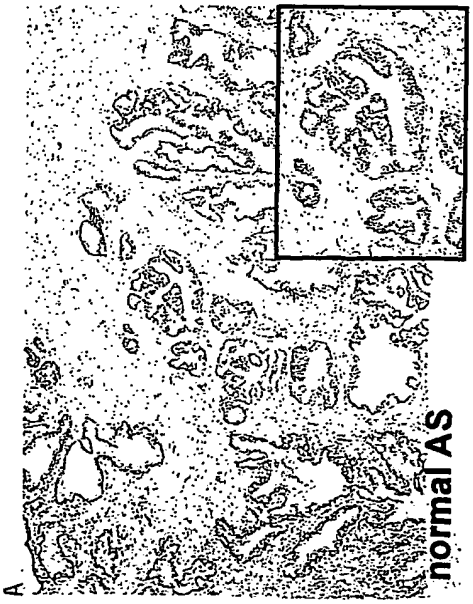

A
b
.5
.5
.4
.4
.35
heart
brain
placenta
lung
liver
skeletal muscle
kidney
pancreas
spleen
thymus
prostate
testis
ovary
small intestine
colon
leukocytes
prostate *
placenta

Fig. 5A

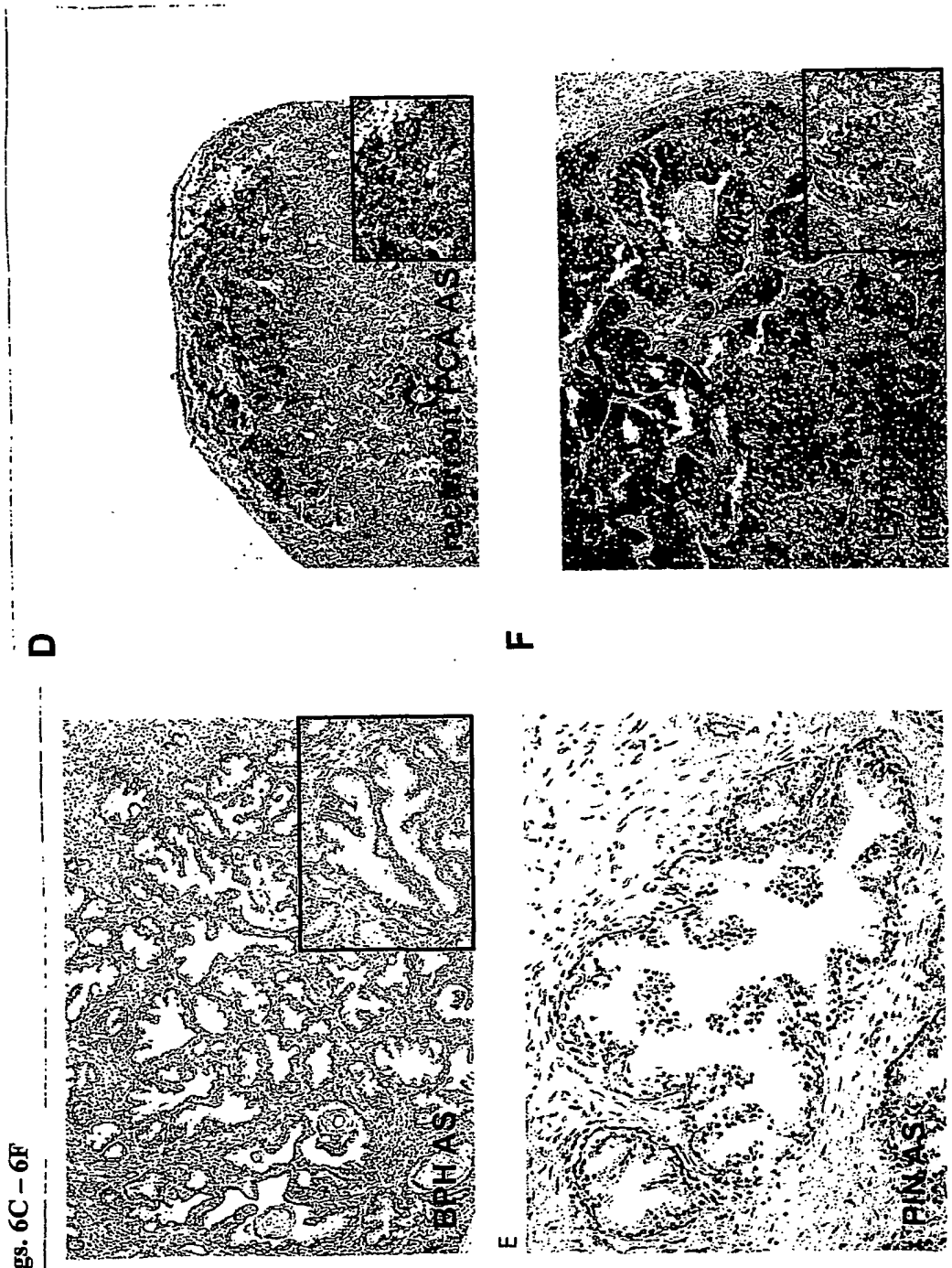
Figs. 5B – 5D







Figs. 6A and 6B



Figs. 6C – 6F

Fig. 7

```
        10                  30                  50
GCCAAGTGTAACAAACTCACAGCCCTCTCCAAACTGGCTGGGGCTGCTGGGAGACTCCCA
        70                  90                 110
AGGAACTCGTCAGGAAGGCAGGAGACAGGAGACGGGACCTCTACAGGGAGACGGTGGGCC
       130                 150                 170
GGCCCTTGGGGGGGCTGATGTGGCCCCAAGGCTGAGTCCCGTCAGGGTCTGGCCTCGGCC
       190                 210                 230
TCAGGCCCCCAAGGAGCCGGCCCTACACCCCATGGGTTTGTCACTGCCCAAGGAGAAAGG
                                  M   G   L   S   L   P   K   E   K   G
       250                 270                 290
GCTAATTCTCTGCCTATGGAGCAAGTTCTGCAGATGGTTCCAGAGACGGGAGTCCTGGGC
 L   I   L   C   L   W   S   K   F   C   R   W   F   Q   R   R   E   S   W   A
       310                 330                 350
CCAGAGCCGAGATGAGCAGAACCTGCTGCAGCAGAAGAGGATCTGGGAGTCTCCTCTCCT
 Q   S   R   D   E   Q   N   L   L   Q   Q   K   R   I   W   E   S   P   L   L
       370                 390                 410
TCTAGCTGCCAAAGATAATGATGTCCAGGCCCTGAACAAGTTGCTCAAGTATGAGGATTG
 L   A   A   K   D   N   D   V   Q   A   L   N   K   L   L   K   Y   E   D   C
       430                 450                 470
CAAGGTGCACCAGAGAGGAGCCATGGGGGAAACAGCGCTACACATAGCAGCCCTCTATGA
 K   V   H   Q   R   G   A   M   G   E   T   A   L   H   I   A   A   L   Y   D
       490                 510                 530
CAACCTGGAGGCCGCCATGGTGCTGATGGAGGCTGCCCCGGAGCTGGTCTTTGAGCCCAT
 N   L   E   A   A   M   V   L   M   E   A   A   P   E   L   V   F   E   P   M
       550                 570                 590
GACATCTGAGCTCTATGAGGGTCAGACTGCACTGCACATCGCTGTTGTGAACCAGAACAT
 T   S   E   L   Y   E   G   Q   T   A   L   H   I   A   V   V   N   Q   N   M
       610                 630                 650
GAACCTGGTGCGAGCCCTGCTTGCCCGCAGGGCCAGTGTCTCTGCCAGAGCCACAGGCAC
 N   L   V   R   A   L   L   A   R   R   A   S   V   S   A   R   A   T   G   T
       670                 690                 710
TGCCTTCCGCCGTAGTCCCCGCAACCTCATCTACTTTGGGGAGCACCCTTTGTCCTTTGC
 A   F   R   R   S   P   R   N   L   I   Y   F   G   E   H   P   L   S   F   A
       730                 750                 770
TGCCTGTGTGAACAGTGAGGAGATCGTGCGGCTGCTCATTGAGCATGGAGCTGACATCCG
 A   C   V   N   S   E   E   I   V   R   L   L   I   E   H   G   A   D   I   R
       790                 810                 830
GGCCCAGGACTCCCTGGGAAACACAGTGTTACACATCCTCATCCTCCAGCCCAACAAAAC
 A   Q   D   S   L   G   N   T   V   L   H   I   L   I   L   Q   P   N   K   T
       850                 870                 890
CTTTGCCTGCCAGATGTACAACCTGTTGCTGTCCTACGACAGACATGGGGACCACCTGCA
 F   A   C   Q   M   Y   N   L   L   L   S   Y   D   R   H   G   D   H   L   Q
       910                 930                 950
GCCCCTGGACCTCGTGCCCAATCACCAGGGTCTCACCCCTTTCAAGCTGGCTGGAGTGGA
 P   L   D   L   V   P   N   H   Q   G   L   T   P   F   K   L   A   G   V   E
       970                 990                1010
GGGTAACACTGTGATGTTTCAGCACCTGATGCAGAAGCGGAAGCACACCCAGTGGACGTA
 G   N   T   V   M   F   Q   H   L   M   Q   K   R   K   H   T   Q   W   T   Y
      1030                1050                1070
TGGACCACTGACCTCGACTCTCTATGACCTCACAGAGATCGACTCCTCAGGGGATGAGCA
 G   P   L   T   S   T   L   Y   D   L   T   E   I   D   S   S   G   D   E   Q
      1090                1110                1130
GTCCCTGCTGGAACTTATCATCACCACCAAGAAGCGGGAGGCTCGCCAGATCCTGGACCA
 S   L   L   E   L   I   I   T   T   K   K   R   E   A   R   Q   I   L   D   Q
      1150                1170                1190
GACGCCGGTGAAGGAGCTGGTGAGCCTCAAGTGGAAGCGGTACGGGCGGCCGTACTTCTG
 T   P   V   K   E   L   V   S   L   K   W   K   R   Y   G   R   P   Y   F   C
      1210                1230                1250
CATGCTGGGTGCCATATATCTGCTGTACATCATCTGCTTCACCATGTGCTGCATCTACCG
 M   L   G   A   I   Y   L   L   Y   I   I   C   F   T   M   C   C   I   Y   R
      1270                1290                1310
```

```
CCCCCTCAAGCCCAGGACCAATAACCGCACAAGCCCCCGGGACAACACCCTCTTACAGCA
 P  L  K  P  R  T  N  N  R  T  S  P  R  D  N  T  L  L  Q  Q
      1330                1350                1370
GAAGCTACTTCAGGAAGCCTACGTGACCCCTAAGGACGATATCCGGCTGGTCGGGGAGCT
 K  L  L  Q  E  A  Y  V  T  P  K  D  D  I  R  L  V  G  E  L
      1390                1410                1430
GGTGACTGTCATTGGGGCTATCATCATCCTGCTGGTAGAGGTTCCAGACATCTTCAGAAT
 V  T  V  I  G  A  I  I  I  L  L  V  E  V  P  D  I  F  R  M
      1450                1470                1490
GGGGGTCACTCGCTTCTTTGGACAGACCATCCTTGGGGGCCCATTCCATGTCCTCATCAT
 G  V  T  R  F  F  G  Q  T  I  L  G  G  P  F  H  V  L  I  I
      1510                1530                1550
CACCTATGCCTTCATGGTGCTGGTGACCATGGTGATGCGGCTCATCAGTGCCAGCGGGGA
 T  Y  A  F  M  V  L  V  T  M  V  M  R  L  I  S  A  S  G  E
      1570                1590                1610
GGTGGTACCCATGTCCTTTGCACTCGTGCTGGGCTGGTGCAACGTCATGTACTTCGCCCG
 V  V  P  M  S  F  A  L  V  L  G  W  C  N  V  M  Y  F  A  R
      1630                1650                1670
AGGATTCCAGATGCTAGGCCCCTTCACCATCATGATTCAGAAGATGATTTTTGGCGACCT
 G  F  Q  M  L  G  P  F  T  I  M  I  Q  K  M  I  F  G  D  L
      1690                1710                1730
GATGCGATTCTGCTGGCTGATGGCTGTGGTCATCCTGGGCTTTGCTTCAGCCTTCTATAT
 M  R  F  C  W  L  M  A  V  V  I  L  G  F  A  S  A  F  Y  I
      1750                1770                1790
CATCTTCCAGACAGAGGACCCCGAGGAGCTAGGCCACTTCTACGACTACCCCATGGCCCT
 I  F  Q  T  E  D  P  E  E  L  G  H  F  Y  D  Y  P  M  A  L
      1810                1830                1850
GTTCAGCACCTTCGAGCTGTTCCTTACCATCATCGATGGCCCAGCCAACTACAACGTGGA
 F  S  T  F  E  L  F  L  T  I  I  D  G  P  A  N  Y  N  V  D
      1870                1890                1910
CCTGCCCTTCATGTACAGCATCACCTATGCTGCCTTTGCCATCATCGCCACACTGCTCAT
 L  P  F  M  Y  S  I  T  Y  A  A  F  A  I  I  A  T  L  L  M
      1930                1950                1970
GCTCAACCTCCTCATTGCCATGATGGGCGACACTCACTGGCGAGTGGCCCATGAGCGGGA
 L  N  L  L  I  A  M  M  G  D  T  H  W  R  V  A  H  E  R  D
      1990                2010                2030
TGAGCTGTGGAGGGCCCAGATTGTGGCCACCACGGTGATGCTGGAGCGGAAGCTGCCTCG
 E  L  W  R  A  Q  I  V  A  T  T  V  M  L  E  R  K  L  P  R
      2050                2070                2090
CTGCCTGTGGCCTCGCTCCGGGATCTGCGGACGGGAGTATGGCCTGGGGGACCGCTGGTT
 C  L  W  P  R  S  G  I  C  G  R  E  Y  G  L  G  D  R  W  F
      2110                2130                2150
CCTGCGGGTGGAAGACAGGCAAGATCTCAACCGGCAGCGGATCCAACGCTACGCACAGGC
 L  R  V  E  D  R  Q  D  L  N  R  Q  R  I  Q  R  Y  A  Q  A
      2170                2190                2210
CTTCCACACCCGGGGCTCTGAGGATTTGGACAAAGACTCAGTGGAAAAACTAGAGCTGGG
 F  H  T  R  G  S  E  D  L  D  K  D  S  V  E  K  L  E  L  G
      2230                2250                2270
CTGTCCCTTCAGCCCCCACCTGTCCCTTCCTACGCCCTCAGTGTCTCGAAGTACCTCCCG
 C  P  F  S  P  H  L  S  L  P  T  P  S  V  S  R  S  T  S  R
      2290                2310                2330
CAGCAGTGCCAATTGGGAAAGGCTTCGGCAAGGGACCCTGAGGAGAGACCTGCGTGGGAT
 S  S  A  N  W  E  R  L  R  Q  G  T  L  R  R  D  L  R  G  I
      2350                2370                2390
AATCAACAGGGGTCTGGAGGACGGGGAGAGCTGGGAATATCAGATCTGACTGCGTGTTCT
 I  N  R  G  L  E  D  G  E  S  W  E  Y  Q  I
      2410                2430                2450
CACTTCGCTTCCTGGAACTTGCTCTCATTTTCCTGGGTGCATCAAACAAAACAAAAACCA
      2470                2490                2510
AACACCCAGAGGTCTCATCTCCCAGGCCCCAGGGAGAAAGAGGAGTAGCATGAACGCCAA
      2530                2550                2570
GGAATGTACGTTGAGAATCACTGCTCCAGGCCTGCATTACTCCTTCAGCTCTGGGGCAGA
```

Fig. 7 / continuation 2

```
      2590                2610                2630
GGAAGCCCAGCCCAAGCACGGGGCTGGCAGGGCGTGAGGAACTCTCCTGTGGCCTGCTCA
      2650                2670                2690
TCACCCTTCCGACAGGAGCACTGCATGTCAGAGCACTTTAAAAACAGGCCAGCCTGCTTG
      2710                2730                2750
GGCCCTCGGTCTCCACCCCAGGGTCATAAGTGGGGAGAGAGCCCTTCCCAGGGCACCCAG
      2770                2790                2810
GCAGGTGCAGGGAAGTGCAGAGCTTGTGGAAAGCGTGTGAGTGAGGGAGACAGGAACGGC
      2830                2850                2870
TCTGGGGGTGGGAAGTGGGGCTAGGTCTTGCCAACTCCATCTTCAATAAAGTCGTTTTCG
      2890                2910
GATCCCTAAAAAAAAAAAAAAAAAAAAAAAAA
```

```
MGLSLPKEKGLILCLWSKFCRWFQRRESWAQSRDEQNLLQQKRIWESPLLLAAKDNDVQALNKLLKYEDCKVHQRGAMGETALHIA
ALYDNLEAAMVLMEAAPELVFEPMTSELYEGQTALHIAVVNQNMNLVRALLARRASVSARATGTAFRRSPRNLIYFGEHPLSFAAC
VNSEEIVRLLIEHGADIRAQDSLGNTVLHILILQPNKTFACQMYNLLLSYDRHGDHLQPLDLVPNHQGLTPFKLAGVEGNTVMFQH
LMQKRKHTQWTYGPLTSTLYDLTEIDSSGDEQSLLELIITTKKREARQILDQTPVKELVSLKWKRYGRPYFCMLGAIYLLYIICFT
MCCIYRPLKPRTNNRTSPRDNTLLQQKLLQEAYVTPKDDIRLVGELVTVIGAIIILLVEVPDIFRMGVTRFFGQTILGGPFHVLII
TYAFMVLVTMVMRLISASGEVVPMSFALVLGWCNVMYFARGFQMLGPFTIMIQKMIFGDLMRFCWLMAVVILGFASAFYIIFQTED
PEELGHFYDYPMALFSTFELFLTIIDGPANYNVDLPFMYSITYAAFAIIATLLMLNLLIAMMGDTHWRVAHERDELWRAQIVATTV
MLERKLPRCLWPRSGICGREYGLGDRWFLRVEDRQDLNRQRIQRYAQAFHTRGSEDLDKDSVEKLELGCPFSPHLSLPTPSVSRST
SRSSANWERLRQGTLRRDLRGIINRGLEDGESWEYQI
```

Figure 8:

A)

```
                    ATGGGTTTGTCACTGCCCAAGGAGAAAGGGCTAATTCTCT
                    M  G  L  S  L  P  K  E  K  G  L  I  L  C
       250                  270                  290
GCCTATGGAGCAAGTTCTGCAGATGGTTCCAGAGACGGGAGTCCTGGGCCCAGAGCCGAG
  L  W  S  K  F  C  R  W  F  Q  R  R  E  S  W  A  Q  S  R  D
       310                  330                  350
ATGAGCAGAACCTGCTGCAGCAGAAGAGGATCTGGGAGTCTCCTCTCCTTCTAGCTGCCA
  E  Q  N  L  L  Q  Q  K  R  I  W  E  S  P  L  L  L  A  A  K
       370                  390                  410
AAGATAATGATGTCCAGGCCCTGAACAAGTTGCTCAAGTATGAGGATTGCAAGGTGCACC
  D  N  D  V  Q  A  L  N  K  L  L  K  Y  E  D  C  K  V  H  Q
       430                  450                  470
AGAGAGGAGCCATGGGGGAAACAGCGCTACACATAGCAGCCCTCTATGACAACCTGGAGG
  R  G  A  M  G  E  T  A  L  H  I  A  A  L  Y  D  N  L  E  A
       490                  510                  530
CCGCCATGGTGCTGATGGAGGCTGCCCCGGAGCTGGTCTTTGAGCCCATGACATCTGAGC
  A  M  V  L  M  E  A  A  P  E  L  V  F  E  P  M  T  S  E  L
       550                  570                  590
TCTATGAGGGTCAGACTGCACTGCACATCGCTGTTGTGAACCAGAACATGAACCTGGTGC
  Y  E  G  Q  T  A  L  H  I  A  V  V  N  Q  N  M  N  L  V  R
       610                  630                  650
GAGCCCTGCTTGCCCGCAGGGCCAGTGTCTCTGCCAGAGCCACAGGCACTGCCTTCCGCC
  A  L  L  A  R  R  A  S  V  S  A  R  A  T  G  T  A  F  R  R
       670                  690                  710
GTAGTCCCTGCAACCTCATCTACTTTGGGGAGCACCCTTTGTCCTTTGCTGCCTGTGTGA
  S  P  C  N  L  I  Y  F  G  E  H  P  L  S  F  A  A  C  V  N
```

Fig. 8 / continuation 1

```
      730                 750                 770
ACAGTGAGGAGATCGTGCGGCTGCTCATTGAGCATGGAGCTGACATCCGGGCCCAGGACT
  S  E  E  I  V  R  L  L  I  E  H  G  A  D  I  R  A  Q  D  S
      790                 810                 830
CCCTGGGAAACACAGTGTTACACATCCTCATCCTCCAGCCCAACAAAACCTTTGCCTGCC
  L  G  N  T  V  L  H  I  L  I  L  Q  P  N  K  T  F  A  C  Q
      850                 870                 890
AGATGTACAACCTGTTGCTGTCCTACGACAGACATGGGGACCACCTGCAGCCCCTGGACC
  M  Y  N  L  L  L  S  Y  D  R  H  G  D  H  L  Q  P  L  D  L
      910                 930                 950
TCGTGCCCAATCACCAGGGTCTCACCCCTTTCAAGCTGGCTGGAGTGGAGGGTAACACTG
  V  P  N  H  Q  G  L  T  P  F  K  L  A  G  V  E  G  N  T  V
      970                 990                1010
TGATGTTTCAGCACCTGATGCAGAAGCGGAAGCACACCCAGTGGACGTATGGACCACTGA
  M  F  Q  H  L  M  Q  K  R  K  H  T  Q  W  T  Y  G  P  L  T
     1030                1050                1070
CCTCGACTCTCTATGACCTCACAGAGATCGACTCCTCAGGGGATGAGCAGTCCCTGCTGG
  S  T  L  Y  D  L  T  E  I  D  S  S  G  D  E  Q  S  L  L  E
     1090                1110                1130
AACTTATCATCACCACCAAGAAGCGGGAGGCTCGCCAGATCCTGGACCAGACGCCGGTGA
  L  I  I  T  T  K  K  R  E  A  R  Q  I  L  D  Q  T  P  V  K
     1150                1170                1190
AGGAGCTGGTGAGCCTCAAGTGGAAGCGGTACGGGCGGCCGTACTTCTGCATGCTGGGTG
  E  L  V  S  L  K  W  K  R  Y  G  R  P  Y  F  C  M  L  G  A
     1210                1230                1250
CCATATATCTGCTGTACATCATCTGCTTCACCATGTGCTGCATCTACCGCCCCCTCAAGC
  I  Y  L  L  Y  I  I  C  F  T  M  C  C  I  Y  R  P  L  K  P
     1270                1290                1310
CCAGGACCAATAACCGCACGAGCCCCCGGGACAACACCCTCTTACAGCAGAAGCTACTTC
  R  T  N  N  R  T  S  P  R  D  N  T  L  L  Q  Q  K  L  L  Q
     1330                1350                1370
AGGAAGCCTACATGACCCCTAAGGACGATATCCGGCTGGTCGGGGAGCTGGTGACTGTCA
  E  A  Y  M  T  P  K  D  D  I  R  L  V  G  E  L  V  T  V  I
     1390                1410                1430
TTGGGGCTATCATCATCCTGCTGGTAGAGGTTCCAGACATCTTCAGAATGGGGGTCACTC
  G  A  I  I  I  L  L  V  E  V  P  D  I  F  R  M  G  V  T  R
     1450                1470                1490
GCTTCTTTGGACAGACCATCCTTGGGGGCCCATTCCATGTCCTCATCATCACCTATGCCT
  F  F  G  Q  T  I  L  G  G  P  F  H  V  L  I  I  T  Y  A  F
     1510                1530                1550
TCATGGTGCTGGTGACCATGGTGATGCGGCTCATCAGTGCCAGCGGGGAGGTGGTACCCA
  M  V  L  V  T  M  V  M  R  L  I  S  A  S  G  E  V  V  P  M
     1570                1590                1610
TGTCCTTTGCACTCGTGCTGGGCTGGTGCAACGTCATGTACTTCGCCCGAGGATTCCAGA
  S  F  A  L  V  L  G  W  C  N  V  M  Y  F  A  R  G  F  Q  M
     1630                1650                1670
TGCTAGGCCCCTTCACCATCATGATTCAGAAGATGATTTTTGGCGACCTGATGCGATTCT
  L  G  P  F  T  I  M  I  Q  K  M  I  F  G  D  L  M  R  F  C
     1690                1710                1730
GCTGGCTGATGGCTGTGGTCATCCTGGGCTTTGCTTCAGCCTTCTATATCATCTTCCAGA
  W  L  M  A  V  V  I  L  G  F  A  S  A  F  Y  I  I  F  Q  T
     1750                1770                1790
CAGAGGACCCCGAGGAGCTAGGCCACTTCTACGACTACCCCATGGCCCTGTTCAGCACCT
  E  D  P  E  E  L  G  H  F  Y  D  Y  P  M  A  L  F  S  T  F
     1810                1830                1850
TCGAGCTGTTCCTTACCATCATCGATGGCCCAGCCAACTACAACGTGGACCTGCCCTTCA
  E  L  F  L  T  I  I  D  G  P  A  N  Y  N  V  D  L  P  F  M
     1870                1890                1910
TGTACAGCATCACCTATGCTGCCTTTGCCATCATCGCCACACTGCTCATGCTCAACCTCC
  Y  S  I  T  Y  A  A  F  A  I  I  A  T  L  L  M  L  N  L  L
     1930                1950                1970
TCATTGCCATGATGGGCGACACTCACTGGCGAGTGGCCCATGAGCGGGATGAGCTGTGGA
```

Fig. 8 / continu on 2

```
  I  A  M  M  G  D  T  H  W  R  V  A  H  E  R  D  E  L  W  R
     1990              2010              2030
GGGCCCAGATTGTGGCCACCACGGTGATGCTGGAGCGGAAGCTGCCTCGCTGCCTGTGGC
  A  Q  I  V  A  T  T  V  M  L  E  R  K  L  P  R  C  L  W  P
     2050              2070              2090
CTCGCTCCGGGATCTGCGGACGGGAGTATGGCCTGGGAGACCGCTGGTTCCTGCGGGTGG
  R  S  G  I  C  G  R  E  Y  G  L  G  D  R  W  F  L  R  V  E
     2110              2130              2150
AAGACAGGCAAGATCTCAACCGGCAGCGGATCCAACGCTACGCACAGGCCTTCCACACCC
  D  R  Q  D  L  N  R  Q  R  I  Q  R  Y  A  Q  A  F  H  T  R
     2170              2190              2210
GGGGCTCTGAGGATTTGGACAAAGACTCAGTGGAAAAACTAGAGCTGGGCTGTCCCTTCA
  G  S  E  D  L  D  K  D  S  V  E  K  L  E  L  G  C  P  F  S
     2230              2250              2270
GCCCCCACCTGTCCCTTCCTATGCCCTCAGTGTCTCGAAGTACCTCCCGCAGCAGTGCCA
  P  H  L  S  L  P  M  P  S  V  S  R  S  T  S  R  S  S  A  N
     2290              2310              2330
ATTGGGAAAGGCTTCGGCAAGGGACCCTGAGGAGAGACCTGCGTGGGATAATCAACAGGG
  W  E  R  L  R  Q  G  T  L  R  R  D  L  R  G  I  I  N  R  G
     2350              2370              2390
GTCTGGAGGACGGGGAGAGCTGGGAATATCAGATCTGA
  L  E  D  G  E  S  W  E  Y  Q  I  *
```

```
MGLSLPKEKGLILCLWSKFCRWFQRRESWAQSRDEQNLLQQKRIWESPLLLAAKDNDVQALNKLLKYEDCKVHQRGAMGETALHIA
ALYDNLEAAMVLMEAAPELVFEPMTSELYEGQTALHIAVVNQNMNLVRALLARRASVSARATGTAFRRSPCNLIYFGEHPLSFAAC
VNSEEIVRLLIEHGADIRAQDSLGNTVLHILILQPNKTFACQMYNLLLSYDRHGDHLQPLDLVPNHQGLTPFKLAGVEGNTVMFQH
LMQKRKHTQWTYGPLTSTLYDLTEIDSSGDEQSLLELIITTKKRRARQILDQTPVKELVSLKWKRYGRPYFCMLGAIYLLYIICFT
MCCIYRPLKPRTNNRTSPRDNTLLQQKLLQEAYMTPKDDIRLVGELVTVIGAIIILLVEVPDIFRMGVTRFFGQTILGGPFHVLII
TYAFMVLVTMVMRLISASGEVVPMSFALVLGWCNVMYFARGFQMLGPFTIMIQKMIFGDLMRFCWLMAVVILGFASAFYIIFQTED
PEELGHFYDYPMALFSTFELFLTIIDGPANYNVDLPFMYSITYAAFAIIATLLMLNLLIAMMGDTHWRVAHERDELWRAQIVATTV
MLERKLPRCLWPRSGICGREYGLGDRWFLRVEDRQDLNRQRIQRYAQAFHTRGSEDLDKDSVEKLELGCPFSPHLSLPMPSVSRST
SRSSANWERLRQGTLRRDLRGIINRGLEDGESWEYQI
```

B)

```
CAAACTCACAGCCCTCTCCAAACTGGCTGGGGCTGCTGGGAGACTCCCAAGGAACTCGTCAGGAAGGCAGGAGACAGGAGACGGGA
CCTCTACAGGGAGACGGTGGGCCGGCCCTTGGGGGGGCTGATGTGGCCCCAAGGCTGAGTCCCGTCAGGGTCTGGCCTCGGCCTCA
GGCCCCCAAGGAGCCGGCCCTACACCCATGGGTTTGTCACTGCCCAAGGAGAAAGGGCTAATTCTCTGCCTATGGAGCAAGTTCT
GCAGATGGTTCCAGAGACGGGAGTCCTGGGCCCAGAGCCGAGATGAGCAGAACCTGCTGCAGCAGAAGAGGATCTGGGAGTCTCCT
CTCCTTCTAGCTGCCAAAGATAATGATGTCCAGGCCCTGAACAAGTTGCTCAAGTATGAGGATTGCAAGGTGCACCAGAGAGGAGC
CATGGGGGAAACAGCGCTACACATAGCAGCCCTCTATGACAACCTGGAGGCCGCCATGGTGCTGATGGAGGCTGCCCCGGAGCTGG
TCTTTGAGCCCATGACATCTGAGCTCTATGAGGGTCAGACTGCACTGCACATCGCTGTTGTGAACCAGAACATGAACCTGGTGCGA
GCCCTGCTTGCCCGCAGGGCCAGTGTCTCTGCCAGAGCCACAGGCACTGCCTTCCGCCGTAGTCCCCGCAACCTCATCTACTTTGG
GGAGCACCCTTTGTCCTTTGCTGCCTGTGTGAACAGTGAGGAGATCGTGCGGCTGCTCATTGAGCATGGAGCTGACATCCGGGCCC
AGGACTGCCTGGCCCAACAAAACCTTTGCCTGCCAGATGTACAACCTGTTGCTGTCCTACGACAGACATGGGGACCACCTGCAGCC
CCTGGACCTCGTGCCCAATCACCAGGGTCTCACCCCTTTCAAGCTGGCTGGAGTGGAGGGTAACACTGTGATGTTTCAGCACCTGA
TGCAGAAGCGGAAGCACACCCAGTGGACGTATGGACCACTGACCTCGACTCTCTATGACCTCACAGAGATCGACTCCTCAGGGGAT
GAGCAGTCCCTGCTGGAACTTATCATCACCACCAAGAAGCGGGAGGCTCGCCAGATCCTGGACCAGACGCCGGTGAAGGAGCTGGT
GAGCCTCAAGTGGAAGCGGTACGGGCGGCCGTACTTCTGCATGCTGGGTGCCATATATCTGCTGTACATCATCTGCTTCACCATGT
GCTGCATCTACCGCCCCCTCAAGCCCAGGACCAATAACCGCACAAGCCCCCGGGACAACACCCTCTTACAGCAGAAGCTACTTCAG
GAAGCCTACGTGACCCCTAAGGACGATATCCGGCTGGTCGGGGAGCTGGTGACTGTCATTGGGGCTATCATCATCCTGCTGGTAGA
GGTTCCAGACATCTTCAGAATGGGGGTCACTCGCTTCTTTGGACAGACCATCCTTGGGGGCCCATTCCATGTCCTCATCATCACCT
ATGCCTTCATGGTGCTGGTGACCATGGTGATGCGGCTCATATGATTTTTGGCGACCTGATGCGATTCTGCTGGCTGATGGCTGTGG
TCATCCTGGGCTTTGCTTCAGCCTTCTATATCATCTTCCAGACAGAGGACCCCGAGGAGCTAGGCCACTTCTACGACTACCCCATG
GCCCTGTTCAGCACCTTCGAGCTGGTCCTTACCATCATCGATGGCCCAGCCAACTACAACGTGGACCTGCCCTTCATGTACAGCAT
CACCTATGCTGCCTTTGCCATCATCGCCACACTGCTCATGCTCAACCTCCTCATTGCCATGATGGGCGACACTCACTGGCGAGTGG
CCCATGAGCGGGATGAGCTGTGGAGGGCCCAGATTGTGGCCACCACGGTGATGCTGGAGCGGAAGCTGCCTCGCTGCCTGTGGCCT
CGCTCCGGGATCTGCGGACGGGAGTATGGCCTGGGGGACCGCTGGTTCCTGCGGGTGGAAGACAGGCAAGATCTCAACCGGCAGCG
```

GATCCAACGCTACGCACAGGCCTTCCACACCCGGGGCTCTGAGGATTTGGACAAAGACTCAGTGGAAAAACTAGAGCTGGGCTGTC
CCTTCAGCCCCCACCTGTCCCTTCCTACGCCCTCAGTGTCTCGAAGTACCTCCCGCAGCAGTGCCAATTGGGAAAGGCTTCGGCAA
GGGACCCTGAGGAGAGACCTGCGTGGGATAATCAACAGGGGTCTGGAGGACGGGGAGAGCTGGGAATATCAGATCTGACTGCGTGT
TCTCACTTCGCTTCCTGGAACTTGCTCTCATTTTCCTGGGTGCATCAAACAAAACAAAAACCAAACACCCAGAGGTCTCATCTCCC
AGGCCCCAGGGAGAAAGAGGAGTAGCATGAACGCCAAGGAATGTACGTTGAGAATCACTGCTCCAGGCCTGCATTACTCCTTCAGC
TCTGGGGCAGAGGAAGCCCAGCCCAAGCACGGGGCTGGCAGGGCGTGAGGAACTCTCCTGTGGCCTGCTCATCACCCTTCCGACAG
GAGCACTGCATGTCAGAGCACTTTAAAAACAGGCCAGCCTGCTTGGGCCCTCGGTCTCCACCCCAGGGTCATAAGTGGGGAGAGAG
CCCTTCCCAGGGCACCCAGGCAGGTGCAGGGAAGTGCAGAGCTTGTGGAAAGCGTGTGAGTGAGGGAGACAGGAACGGCTCTGGGG
GTGGGAAGTGGGGCTAGGTCTTGCCAACTCCATCTTCAATAAAGTCGTTTTCGGATCCCTAAAAAAAAAAAAAAAAAAAAAAAAAA

c.)

CAAACTCACAGCCCTCTCCAAACTGGCTGGGGCTGCTGGGAGACTCCCAAGGAACTCGTCAGGAAGGCAGGAGACAGGAGACGGGA
CCTCTACAGGGAGACGGTGGGCCGGCCCTTGGGGGGGCTGATGTGGCCCCAAGGCTGAGTCCCGTCAGGGTCTGGCCTCGGCCTCA
GGCCCCCAAGGAGCCGGCCCTACACCCCATGGGTTTGTCACTGCCCAAGGAGAAAGGGCTAATTCTCTGCCTATGGAGCAAGTTCT
GCAGATGGTTCCAGAGACGGGAGTCCTGGGCCCAGAGCCGAGATGAGCAGAACCTGCTGCAGCAGAAGAGGATCTGGGAGTCTCCT
CTCCTTCTAGCTGCCAAAGATAATGATGTCCAGGCCCTGAACAAGTTGCTCAAGTATGAGGATTGCAAGGTGCACCAGAGAGGAGC
CATGGGGGAAACAGCGCTACACATAGCAGCCCTCTATGACAACCTGGAGGCCGCCATGGTGCTGATGGAGGCTGCCCCGGAGCTGG
TCTTTGAGCCCATGACATCTGAGCTCTATGAGGTCCTGACTGCCCATCACTTGAACGCCTGCCCCCTGAAATGCCAGGGCCTAGAG
AAGAGGAAGAGATGGGCAGCAGCTGGATCCCCTGGGAATCCTGAACACCCGAGAGCTCCCTGTTCTCCATCCCAGGCTACCCCTGA
GGGAAAGAGACTGGGGTGCATATGGGAGGGACCCCCTGCAGGATCCTGGGGACAGACCCGTGACTGACAGCTGTCTCTGGGCCAGG
TCAGACTGCACTGCACATCGCTGTTGTGAACCAGAACATGAACCTGGTGCGAGCCCTGCTTGCCCGCAGGGCCAGTGTCTCTGCCA
GAGCCACAGGCACTGCCTTCCGCCGTAGTCCCTGCAACCTCATCTACTTTGGGGAGCACCCTTTGTCCTTTGCTGCCTGTGTGAAC
AGTGAGGAGATCGTGCGGCTGCTCATTGAGCATGGAGCTGACATCCGGGCCCAGGACTCCCTGGCCCAACAAAACCTTTGCCTGCC
AGATGTACAACCTGTTGCTGTCCTACGACAGACATGGGGACCACCTGCAGCCCCTGGACCTCGTGCCCAATCACCAGGGTCTCACC
CCTTTCAAGCTGGCTGGAGTGGAGGGTAACACTGTGATGTTTCAGCACCTGATGCAGAAGCGGAAGCACACCCAGTGGACGTATGG
ACCACTGACCTCGACTCTCTATGACCTCACAGAGATCGACTCCTCAGGGGATGAGCAGTCCCTGCTGGAACTTATCATCACCACCA
AGAAGCGGGAGGCTCGCCAGATCCTGGACCAGACGCCGGTGAAGGAGCTGGTGAGCCTCAAGTGGAAGCGGTACGGGCGGCCGTAC
TTCTGCATGCTGGGTGCCATATATCTGCTGTACATCATCTGCTTCACCATGTGCTGCATCTACCGCCCCCTCAAGCCCAGGACCAA
TAACCGCACGAGCCCCCGGGACAACACCCTCTTACAGCAGAAGCTACTTCAGGAAGCCTACATGACCCCTAAGGACGATATCCGGC
TGGTCGGGGAGCTGGTGACTGTCATTGGGGCTATCATCATCCTGCTGGTAGAGGTTCCAGACATCTTCAGAATGGGGGTCACTCGC
TTCTTTGGACAGACCATCCTTGGGGGCCCATTCCATGTCCTCATCATCACCTATGCCTTCATGGTGCTGGTGACCATGGTGATGCG
GCTCATCAGTGCCAGCGGGGAGGTGGTACCCATGTCCTTTGCACTCGTGCTGGGCTGGTGCAACGTCATGTACTTCGCCCGAGGAT
TCCAGATGCTAGGCCCCTTCACCATCATGATTCAGAAGATGATTTTTGGCGACCTGATGCGATTCTGCTGGCTGATGGCTGTGGTC
ATCCTGGGCTTTGCTTAGACAGAGGACCCCGAGGAGCTAGGCCACTTCTACGACTACCCCATGGCCCTGTTCAGCACCTTCGAGCT
GGTCCTTACCATCATCGATGGCCCAGCCAACTACAACGTGGACCTGCCCTTCATGTACAGCATCACCTATGCTGCCTTTGCCATCA
TCGCCACACTGCTCATGCTCAACCTCCTCATTGCCATGATGGGCGACACTCACTGGCGAGTGGCCCATGAGCGGGATGAGCTGTGG
AGGGCCCAGATTGTGGCCACCACGGTGATGCTGGAGCGGAAGCTGCCTCGCTGCCTGTGGCCTCGCTCCGGGATCTGCGGACGGGA
GTATGGCCTGGGAGACCGCTGGTTCCTGCGGGTGGAAGACAGGCAAGATCTCAACCGGCAGCGGATCCAACGCTACGCACAGGCCT
TCCACACCCGGGGCTCTGAGGATTTGGACAAAGACTCAGTGGAAAAACTAGAGCTGGGCTGTCCCTTCAGCCCCCACCTGTCCCTT
CCTATGCCCTCAGTGTCTCGAAGTACCTCCCGCAGCAGTGCCAATTGGGAAAGGCTTCGGCAAGGGACCCTGAGGAGAGACCTGCG
TGGGATAATCAACAGGGGTCTGGAGGACGGGGAGAGCTGGGAATATCAGATCTGACTGCGTGTTCTCACTTCGCTTCCTGGAACTT
GCTCTCATTTTCCTGGGTGCATCAAACAAAACAAAAACCAAACACCCAGAGGTCTCATCTCCCAGGCCCCAGGGAGAAAGAGGAGT
AGCATGAACGCCAAGGAATGTACGTTGAGAATCACTGCTCCAGGCCTGCATTACTCCTTCAGCTCTGGGGCAGAGGAAGCCCAGCC
CAAGCACGGGGCTGGCAGGGCGTGAGGAACTCTCCTGTGGCCTGCTCATCACCCTTCCGACAGGAGCACTGCATGTCAGAGCACTT
TAAAAACAGGCCAGCCTGCTTGGGCCCTCGGTCTCCACCCCAGGGTCATAAGTGGGGAGAGAGCCCTTCCCAGGGCACCCAGGCAG
GTGCAGGGAAGTGCAGAGCTTGTGGAAAGCGTGTGAGTGAGGGAGACAGGAACGGCTCTGGGGGTGGGAAGTGGGGCTAGGTCTTG
CCAACTCCATCTTCAATAAAGTCGTTTTCGGATCCCTAAAAAAAAAAAAAAAAAAAAAAAAAA

D.)

CAAACTCACAGCCCTCTCCAAACTGGCTGGGGCTGCTGGGAGACTCCCAAGGAACTCGTCAGGAAGGCAGGAGACAGGAGACGGGA
CCTCTACAGGGAGACGGTGGGCCGGCCCTTGGGGGGGCTGATGTGGCCCCAAGGCTGAGTCCCGTCAGGGTCTGGCCTCGGCCTCA
GGCCCCCAAGGAGCCGGCCCTACACCCCATGGGTTTGTCACTGCCCAAGGAGAAAGGGCTAATTCTCTGCCTATGGAGCAAGTTCT
GCAGATGGTTCCAGAGACGGGAGTCCTGGGCCCAGAGCCGAGATGAGCAGAACCTGCTGCAGCAGAAGAGGATCTGGGAGTCTCCT
CTCCTTCTAGCTGCCAAAGATAATGATGTCCAGGCCCTGAACAAGTTGCTCAAGTATGAGGATTGCAAGGTGCACCAGAGAGGAGC
CATGGGGGAAACAGCGCTACACATAGCAGCCCTCTATGACAACCTGGAGGCCGCCATGGTGCTGATGGAGGCTGCCCCGGAGCTGG
TCTTTGAGCCCATGACATCTGAGCTCTATGAGGGTCAGACTGCACTGCACATCGCTGTTGTGAACCAGAACATGAACCTGGTGCGA
GCCCTGCTTGCCCGCAGGGCCAGTGTCTCTGCCAGAGCCACAGGCACTGCCTTCCGCCGTAGTCCCCGCAACCTCATCTACTTTGG

AAACACAGTGTTACACATCCTCATCCTCCAGCCCAACAAAACCTTTGCCTGCCAGATGTACAACCTGTTGCTGTCCTACGACAGAC
ATGGGGACCACCTGCAGCCCCTGGACCTCGTGCCCAATCACCAGGGTCTCACCCCTTTCAAGCTGGCTGGAGTGGAGGGTAACACT
GTGATGTTTCAGCACCTGATGCAGAAGCGGAAGCACACCCAGTGGACGTATGGACCACTGACCTCGACTCTCTATGACCTCACAGA
GATCGACTCCTCAGGGGATGAGCAGTCCCTGCTGGAACTTATCATCACCACCAAGAAGCGGGAGGCTCGCCAGATCCTGGACCAGA
CGCCGGTGAAGGAGCTGGTGAGCCTCAAGTGGAAGCGGTACGGGCGGCCGTACTTCTGCATGCTGGGTGCCATATATCTGCTGTAC
ATCATCTGCTTCACCATGTGCTGCATCTACCGCCCCCTCAAGCCCAGGACCAATAACCGCACAAGCCCCCGGGACAACACCCTCTT
ACAGCAGAAGCTACTTCAGGAAGCCTACGTGACCCCTAAGGACGATATCCGGCTGGTCGGGGAGCTGGTGACTGTCATTGGGGCTA
TCATCATCCTGCTGGTAGAGGTTCCAGACATCTTCAGAATGGGGGTCACTCGCTTCTTTGGACAGACCATCCTTGGGGGCCCATTC
CATGTCCTCATCATCACCTATGCCTTCATGGTGCTGGTGACCATGGTGATGCGGCTCATCAGTGCCAGCGGGGAGGTGGTACCCAT
GTCCTTTGCACTCGTGCTGGGCTGGTGCAACGTCATGTACTTCGCCCGAGGATTCCAGATGCTAGGCCCCTTCACCATCATGATTC
AGAAGATGATTTTTGGCGACCTGATGCGATTCTGCTGGCTGATGGCTGTGGTCATCCTGGGCTTTGCTTCAGCCTTCTATATCATC
TTCCAGACAGAGGACCCCGAGGAGCTAGGCCACTTCTACGACTACCCCATGGCCCTGTTCAGCACCTTCGAGCTGGTCCTTACCAT
CATCGATGGCCCAGCCAACTACAACGTGGACCTGCCCTTCATGTACAGCATCACCTATGCTGCCTTTGCCATCATCGCCACACTGC
TCATGCTCAACCTCCTCATTGCCATGATGGGCGACACTCACTGGCGAGTGGCCCATGAGCGGGATGAGCTGTGGAGGGCCCAGATT
GTGGCCACCACGGTGATGCTGGAGCGGAAGCTGCCTCGCTGCCTGTGGCCTCGCTCCGGGATCTGCGGACGGGAGTATGGCCTGGG
GGACCGCTGGTTCCTGCGGGTGGAAGACAGGCAAGATCTCAACCGGCAGCGGATCCAACGCTACGCACAGGCCTTCCACACCCGGG
GCTCTGAGGATTTGGACAAAGACTCAGTGGAAAAACTAGAGCTGGGCTGTCCCTTCAGCCCCCACCTGTCCCTTCCTACGCCCTCA
GTGTCTCGAAGTACCTCCCGCAGCAGTGCCAATTGGGAAAGGCTTCGGCAAGGGACCCTGAGGAGAGACCTGCGTGGGATAATCAA
CAGGGGTCTGGAGGACGGGGAGAGCTGGGAATATCAGATCTGACTGCGTGTTCTCACTTCGCTTCCTGGAACTTGCTCTCATTTTC
CTGGGTGCATCAAACAAAACAAAAACCAAACACCCAGAGGTCTCATCTCCCAGGCCCCAGGGAGAAAGAGGAGTAGCATGAACGCC
AAGGAATGTACGTTGAGAATCACTGCTCCAGGCCTGCATTACTCCTTCAGCTCTGGGGCAGAGGAAGCCCAGCCCAAGCACGGGGC
TGGCAGGGCGTGAGGAACTCTCCTGTGGCCTGCTCATCACCCTTCCGACAGGAGCACTGCATGTCAGAGCACTTTAAAAACAGGCC
AGCCTGCTTGGGCCCTCGGTCTCCACCCCAGGGTCATAAGTGGGGAGAGAGCCCTTCCCAGGGCACCCAGGCAGGTGCAGGGAAGT
GCAGAGCTTGTGGAAAGCGTGTGAGTGAGGGAGACAGGAACGGCTCTGGGGGTGGGAAGTGGGGCTAGGTCTTGCCAACTCCATCT
TCAATAAAGTCGTTTTCGGATCCCTAAAAAAAAAAAAAAAAAAAAAAAAA

B.)

CACACATGGGGCCTCCCAGGAGTGCCCAGGACCTCGTGCTGTTGGCCTCTGAATCTATCGTCTCCAATCCGCTGTCCCACAGAAGC
CATATAACCCACCTCTCTGTAAATGCCAGGAGCCATGGGGGAAACAGCGCTACACATAGCAGCCCTCTATGACAACCTGGAGGCCG
CCATGGTGCTGATGGAGGCTGCCCCGGAGCTGGTCTTTGAGCCCATGACATCTGAGCTCTATGGAGGGTGAGGGCCCACGGGTCTG
GGGTGAAGAGCAGGAGTGACGTGGTTGGGTATTCAAGTCAGTCTCTGTGATGGATAATTTGGGAAAGACACAGGGGATCTGAGCCT
CCTACTCTTTTTSTCTTCTCTGTCTCCCTTCCGTGTCAGTCCCTGACTGCCCATCACTTGAACGCCTGCCCCCTGAAATGCCAGGG
GCCTAGAGAAGAGGAAGAGATGGGCAGCAGCTGGATCCCCTGGGAATCCTGAACACCCGAGAGCTCCCTGTTCTCCATCCCAGGCT
ACCCCTGAGGGAAAGAGACTAGGGGTGCATATGGGAGGGACCCCCTGCAGGATCCTAGGGGACAGACCCGTGACTGACAGCTGTCT
CTGGGCCAGGTCAGACTGCACTGCACATCGCTGTTGTGAACCAGAACATGAACCTGGTGCGAGCCCTGCTTGCCCGCAGGGCCAGT
GTCTCTGCCAGAGCCACAGGCACTGCCTTCCGCCGTAGTCCCTGCAACCTCATCTACTTTGGGGAGCACCCTTTGTCCTTTGCTGC
CTGTGTGAACAGTGAGGAGATCGTGCGGCTGCTCATTGAGCATGGAGCTGACATCCGGGCCCAGGACTCCCTGGATGTACAACCTG
TTGCTGTCCTACGACAGACATGGGGACCACCTGCAGCCCCTGGACCTCGTGCCCAATCACCAGGGTCTCACCCCTTTCAAGCTGGC
TGGAGTGGAGGGTAACACTGTGATGTTTCAGCACCTGATGCAGAAGCGGAAGCACACCCAGTGGACGTATGGACCACTGACCTCGA
CTCTCTATGACCTCACAGAGATCGACTCCTCAGGGGATGAGCAGTCCCTGCTGGAACTTATCATCACCACCAAGAAGCGGGAGGCT
CGCCAGATCCTGGACCAGACGCCGGTGAAGGAGCTGGTGAGCCTCAAGTGGAAGCGGTACGGGCGGCCGTACTTCTGCATGCTGGG
TGCCATATATCTGCTGTACATCATCTGCTTCACCATGTGCTGCATCTACCGCCCCCTCAAGCCCAGGACCAATAACCGCACGAGCC
CCCGGGACAACACCCTCTTACAGCAGAAGCTACTTCAGGAAGCCTACATGACCCCTAAGGACGATATCCGGCTGGTCGGGGAGCTG
GTGACTGTCATTGGGGCTATCATCATCCTGCTGGTAGAGGTTCCAGACATCTTCAGAATGGGGGTCACTCGCTTCTTTGGACAGAC
CATCCTTGGGGGCCCATTCCATGTCCTCATCATCACCTATGCCTTCATGGTGCTGGTGACCATGGTGATGCGGCTCATCAGTGCCA
GCGGGGAGGTGGTACCCATGTCCTTTGCACTCGTGCTGGGCTGGTGCAACGTCATGTACTTCGCCCGAGGATTCCAGATGCTAGGC
CCCTTCACCATCATGATTCAGAAGATGATTTTTGGCGACCTGATGCGATTCTGCTGGCTGATGGCTGTGGTCATCCTGGGCTTTGC
TTCAGCCTTCTATATCATCTTCCAGACAGAGGACCCCGAGGAGCTAGGCCACTTCTACGACTACCCCATGGCCCTGTTCAGCACCT
TCGAGCTGGTCCTTACCATCATCGATGGCCCAGCCAACTACAACGTGGACCTGCCCTTCATGTACAGCATCACCTATGCTGCCTTT
GCCATCATCGCCACACTGCTCATGCTCAACCTCCTCATTGCCATGATGGGCGACACTCACTGGCGAGTGGCCCATGAGCGGGATGA
GCTGTGGAGGGCCCAGATTGTGGCCACCACGGTGATGCTGGAGCGGAAGCTGCCTCGCTGCCTGTGGCCTCGCTCCGGGATCTGCG
GACGGGAGTATGGCCTGGGAGACCGCTGGTTCCTGCGGGTGGAAGACAGGCAAGATCTCAACCGGCAGCGGATCCAACGCTACGCA
CAGGCCTTCCACACCCGGGGCTCTGAGGATTTGGACAAAGACTCAGTGGAAAAACTAGAGCTGGGCTGTCCCTTCAGCCCCCACCT
GTCCCTTCCTATGCCCTCAGTGTCTCGAAGTACCTCCCGCAGCAGTGCCAATTGGGAAAGGCTTCGGCAAGGGACCCTGAGGAGAG
ACCTGCGTGGGATAATCAACAGGGGTCTGGAGGACGGGGAGAGCTGGGAATATCAGATCTGACTGCGTGTTCTCACTTCGCTTCCT
GGAACTTGCTCTCATTTTCCTGGGTGCATCAAACAAAACAAAAACCAAACACCCAGAGGTCTCATCTCCCAGGCCCCAGGGAGAAA
GAGGAGTAGCATGAACGCCAAGGAATGTACGTTGAGAATCACTGCTCCAGGCCTGCATTACTCCTTCAGCTCTGGGGCAGAGGAAG
CCCAGCCCAAGCACGGGGCTGGCAGGGCGTGAGGAACTCTCCTGTGGCCTGCTCATCACCCTTCCGACAGGAGCACTGCATGTCAG
AGCACTTTAAAAACAGGCCAGCCTGCTTGGGCCCTCGGTCTCCACCCCAGGGTCATAAGTGGGGAGAGAGCCCTTCCCAGGGCACC

Fig. 8 / continuation 5

CAGGCAGGTGCAGGGAAGTGCAGAGCTTGTGGAAAGCGTGTGAGTGAGGGAGACAGGAACGGCTCTGGGGGTGGGAAGTGGGGCTA
GGTCTTGCCAACTCCATCTTCAATAAAGTCGTTTTCGGATCCCTAAAAAAAAAAAAAAAAAAAAAAAAAAA

Figure 9:

A.

```
        10                  30                  50
CGGGGCCCTGGGCTGCAGGAGGTTGCGGCGGCCGCGGCAGCATGGTGGTGCCGGAGAAGG
                                              M  V  V  P  E  K  E
        70                  90                 110
AGCAGAGCTGGATCCCCAAGATCTTCAAGAAGAAGACCTGCACGACGTTCATAGTTGACT
  Q  S  W  I  P  K  I  F  K  K  K  T  C  T  T  F  I  V  D  S
       130                 150                 170
CCACAGATCCGGGAGGGACCTTGTGCCAGTGTGGGCGCCCCCGGACCGCCCACCCCGCAG
  T  D  P  G  G  T  L  C  Q  C  G  R  P  R  T  A  H  P  A  V
       190                 210                 230
TGGCCATGGAGGATGCCTTCGGGGCAGCCGTGGTGACCGTGTGGGACAGCGATGCACACA
  A  M  E  D  A  F  G  A  A  V  V  T  V  W  D  S  D  A  H  T
       250                 270                 290
CCACGGAGAAGCCCACCGATGCCTACGGAGAGCTGGACTTCACGGGGGCCGGCCGCAAGC
  T  E  K  P  T  D  A  Y  G  E  L  D  F  T  G  A  G  R  K  H
       310                 330                 350
ACAGCAATTTCCTCCGGCTCTCTGACCGAACGGATCCAGCTGCAGTTTATAGTCTGGTCA
  S  N  F  L  R  L  S  D  R  T  D  P  A  A  V  Y  S  L  V  T
       370                 390                 410
CACGCACATGGGGCTTCCGTGCCCCGAACCTGGTGGTGTCAGTGCTGGGGGGATCGGGGG
  R  T  W  G  F  R  A  P  N  L  V  V  S  V  L  G  G  S  G  G
       430                 450                 470
GCCCCGTCCTCCAGACCTGGCTGCAGGACCTGCTGCGTCGTGGGCTGGTGCGGGCTGCCC
  P  V  L  Q  T  W  L  Q  D  L  L  R  R  G  L  V  R  A  A  Q
       490                 510                 530
AGAGCACAGGAGCCTGGATTGTCACTGGGGGTCTGCACACGGGCATCGGCCGGCATGTTG
  S  T  G  A  W  I  V  T  G  G  L  H  T  G  I  G  R  H  V  G
       550                 570                 590
GTGTGGCTGTACGGGACCATCAGATGGCCAGCACTGGGGGCACCAAGGTGGTGGCCATGG
  V  A  V  R  D  H  Q  M  A  S  T  G  G  T  K  V  V  A  M  G
       610                 630                 650
GTGTGGCCCCCTGGGGTGTGGTCCGGAATAGAGACACCCTCATCAACCCCAAGGGCTCGT
  V  A  P  W  G  V  V  R  N  R  D  T  L  I  N  P  K  G  S  F
       670                 690                 710
TCCCTGCGAGGTACCGGTGGCGCGGTGACCCGGAGGACGGGGTCCAGTTTCCCCTGGACT
  P  A  R  Y  R  W  R  G  D  P  E  D  G  V  Q  F  P  L  D  Y
       730                 750                 770
ACAACTACTCGGCCTTCTTCCTGGTGGACGACGGCACACACGGCTGCCTGGGGGGCGAGA
  N  Y  S  A  F  F  L  V  D  D  G  T  H  G  C  L  G  G  E  N
       790                 810                 830
ACCGCTTCCGCTTGCGCCTGGAGTCCTACATCTCACAGCAGAAGACGGGCGTGGGAGGGA
  R  F  R  L  R  L  E  S  Y  I  S  Q  Q  K  T  G  V  G  G  T
       850                 870                 890
CTGGAATTGACATCCCTGTCCTGCTCCTCCTGATTGATGGTGATGAGAAGATGTTGACGC
  G  I  D  I  P  V  L  L  L  L  I  D  G  D  E  K  M  L  T  R
       910                 930                 950
GAATAGAGAACGCCACCCAGGCTCAGCTCCCATGTCTCCTCGTGGCTGGCTCAGGGGGAG
  I  E  N  A  T  Q  A  Q  L  P  C  L  L  V  A  G  S  G  G  A
       970                 990                1010
CTGCGGACTGCCTGGCGGAGACCCTGGAAGACACTCTGGCCCCAGGGAGTGGGGGAGCCA
  A  D  C  L  A  E  T  L  E  D  T  L  A  P  G  S  G  G  A  R
      1030                1050                1070
GGCAAGGCGAAGCCCGAGATCGAATCAGGCGTTTCTTTCCCAAAGGGGACCTTGAGGTCC
```

Fig. 9 / continu: n 1

```
 Q  G  E  A  R  D  R  I  R  R  F  F  P  K  G  D  L  E  V  L
     1090                1110                1130
TGCAGGCCCAGGTGGAGAGGATTATGACCCGGAAGGAGCTCCTGACAGTCTATTCTTCTG
 Q  A  Q  V  E  R  I  M  T  R  K  E  L  L  T  V  Y  S  S  E
     1150                1170                1190
AGGATGGGTCTGAGGAATTCGAGACCATAGTTTTGAAGGCCCTTGTGAAGGCCTGTGGGA
 D  G  S  E  E  F  E  T  I  V  L  K  A  L  V  K  A  C  G  S
     1210                1230                1250
GCTCGGAGGCCTCAGCCTACCTGGATGAGCTGCGTTTGGCTGTGGCTTGGAACCGCGTGG
 S  E  A  S  A  Y  L  D  E  L  R  L  A  V  A  W  N  R  V  D
     1270                1290                1310
ACATTGCCCAGAGTGAACTCTTTCGGGGGGACATCCAATGGCGGTCCTTCCATCTCGAAG
 I  A  Q  S  E  L  F  R  G  D  I  Q  W  R  S  F  H  L  E  A
     1330                1350                1370
CTTCCCTCATGGACGCCCTGCTGAATGACCGGCCTGAGTTCGTGCGCTTGCTCATTTCCC
 S  L  M  D  A  L  L  N  D  R  P  E  F  V  R  L  L  I  S  H
     1390                1410                1430
ACGGCCTCAGCCTGGGCCACTTCCTGACCCCGATGCGCCTGGCCCAACTCTACAGCGCGG
 G  L  S  L  G  H  F  L  T  P  M  R  L  A  Q  L  Y  S  A  A
     1450                1470                1490
CGCCCTCCAACTCGCTCATCCGCAACCTTTTGGACCAGGCGTCCCACAGCGCAGGCACCA
 P  S  N  S  L  I  R  N  L  L  D  Q  A  S  H  S  A  G  T  K
     1510                1530                1550
AAGCCCCAGCCCTAAAAGGGGGAGCTGCGGAGCTCCGGCCCCCTGACGTGGGGCATGTGC
 A  P  A  L  K  G  G  A  A  E  L  R  P  P  D  V  G  H  V  L
     1570                1590                1610
TGAGGATGCTGCTGGGGAAGATGTGCGCGCCGAGGTACCCCTCCGGGGGCGCCTGGGACC
 R  M  L  L  G  K  M  C  A  P  R  Y  P  S  G  G  A  W  D  P
     1630                1650                1670
CTCACCCAGGCCAGGGCTTCGGGGAGAGCATGTATCTGCTCTCGGACAAGGCCACCTCGC
 H  P  G  Q  G  F  G  E  S  M  Y  L  L  S  D  K  A  T  S  P
     1690                1710                1730
CGCTCTCGCTGGATGCTGGCCTCGGGCAGGCCCCCTGGAGCGACCTGCTTCTTTGGGCAC
 L  S  L  D  A  G  L  G  Q  A  P  W  S  D  L  L  L  W  A  L
     1750                1770                1790
TGTTGCTGAACAGGGCACAGATGGCCATGTACTTCTGGGAGATGGGTTCCAATGCAGTTT
 L  L  N  R  A  Q  M  A  M  Y  F  W  E  M  G  S  N  A  V  S
     1810                1830                1850
CCTCAGCTCTTGGGGCCTGTTTGCTGCTCCGGGTGATGGCACGCCTGGAGCCTGACGCTG
 S  A  L  G  A  C  L  L  L  R  V  M  A  R  L  E  P  D  A  E
     1870                1890                1910
AGGAGGCAGCACGGAGGAAAGACCTGGCGTTCAAGTTTGAGGGGATGGGCGTTGACCTCT
 E  A  A  R  R  K  D  L  A  F  K  F  E  G  M  G  V  D  L  F
     1930                1950                1970
TTGGCGAGTGCTATCGCAGCAGTGAGGTGAGGGCTGCCCGCCTCCTCCTCCGTCGCTGCC
 G  E  C  Y  R  S  S  E  V  R  A  A  R  L  L  L  R  R  C  P
     1990                2010                2030
CGCTCTGGGGGGATGCCACTTGCCTCCAGCTGGCCATGCAAGCTGACGCCCGTGCCTTCT
 L  W  G  D  A  T  C  L  Q  L  A  M  Q  A  D  A  R  A  F  F
     2050                2070                2090
TTGCCCAGGATGGGGTACAGTCTCTGCTGACACAGAAGTGGTGGGGAGATATGGCCAGCA
 A  Q  D  G  V  Q  S  L  L  T  Q  K  W  W  G  D  M  A  S  T
     2110                2130                2150
CTACACCCATCTGGGCCCTGGTTCTCGCCTTCTTTTGCCCTCCACTCATCTACACCCGCC
 T  P  I  W  A  L  V  L  A  F  F  C  P  P  L  I  Y  T  R  L
     2170                2190                2210
TCATCACCTTCAGGAAATCAGAAGAGGAGCCCACACGGGAGGAGCTAGAGTTTGACATGG
 I  T  F  R  K  S  E  E  E  P  T  R  E  E  L  E  F  D  M  D
     2230                2250                2270
ATAGTGTCATTAATGGGGAAGGGCCTGTCGGGACGGCGGACCCAGCCGAGAAGACGCCGC
 S  V  I  N  G  E  G  P  V  G  T  A  D  P  A  E  K  T  P  L
     2290                2310                2330
```

Fig. 9 / continuation 2

```
TGGGGGTCCCGCGCCAGTCGGGCCGTCCGGGTTGCTGCGGGGGCCGCTGCGGGGGGCGCC
  G  V  P  R  Q  S  G  R  P  G  C  C  G  G  R  C  G  G  R  R
     2350                2370                2390
GGTGCCTACGCCGCTGGTTCCACTTCTGGGGCGTGCCGGTGACCATCTTCATGGGCAACG
  C  L  R  R  W  F  H  F  W  G  V  P  V  T  I  F  M  G  N  V
     2410                2430                2450
TGGTCAGCTACCTGCTGTTCCTGCTGCTTTTCTCGCGGGTGCTGCTCGTGGATTTCCAGC
  V  S  Y  L  L  F  L  L  L  F  S  R  V  L  L  V  D  F  Q  P
     2470                2490                2510
CGGCGCCGCCCGGCTCCCTGGAGCTGCTGCTCTATTTCTGGGCTTTCACGCTGCTGTGCG
  A  P  P  G  S  L  E  L  L  L  Y  F  W  A  F  T  L  L  C  E
     2530                2550                2570
AGGAACTGCGCCAGGGCCTGAGCGGAGGCGGGGGCAGCCTCGCCAGCGGGGGCCCCGGGC
  E  L  R  Q  G  L  S  G  G  G  G  S  L  A  S  G  G  P  G  P
     2590                2610                2630
CTGGCCATGCCTCACTGAGCCAGCGCCTGCGCCTCTACCTCGCCGACAGCTGGAACCAGT
  G  H  A  S  L  S  Q  R  L  R  L  Y  L  A  D  S  W  N  Q  C
     2650                2670                2690
GCGACCTAGTGGCTCTCACCTGCTTCCTCCTGGGCGTGGGCTGCCGGCTGACCCCGGGTT
  D  L  V  A  L  T  C  F  L  L  G  V  G  C  R  L  T  P  G  L
     2710                2730                2750
TGTACCACCTGGGCCGCACTGTCCTCTGCATCGACTTCATGGTTTTCACGGTGCGGCTGC
  Y  H  L  G  R  T  V  L  C  I  D  F  M  V  F  T  V  R  L  L
     2770                2790                2810
TTCACATCTTCACGGTCAACAAACAGCTGGGGCCCAAGATCGTCATCGTGAGCAAGATGA
  H  I  F  T  V  N  K  Q  L  G  P  K  I  V  I  V  S  K  M  M
     2830                2850                2870
TGAAGGACGTGTTCTTCTTCCTCTTCTTCCTCGGCGTGTGGCTGGTAGCCTATGGCGTGG
  K  D  V  F  F  F  L  F  F  L  G  V  W  L  V  A  Y  G  V  A
     2890                2910                2930
CCACGGAGGGGCTCCTGAGGCCACGGGACAGTGACTTCCCAAGTATCCTGCGCCGCGTCT
  T  E  G  L  L  R  P  R  D  S  D  F  P  S  I  L  R  R  V  F
     2950                2970                2990
TCTACCGTCCCTACCTGCAGATCTTCGGGCAGATTCCCCAGGAGGACATGGACGTGGCCC
  Y  R  P  Y  L  Q  I  F  G  Q  I  P  Q  E  D  M  D  V  A  L
     3010                3030                3050
TCATGGAGCACAGCAACTGCTCGTCGGAGCCCGGCTTCTGGGCACACCCTCCTGGGGCCC
  M  E  H  S  N  C  S  S  E  P  G  F  W  A  H  P  P  G  A  Q
     3070                3090                3110
AGGCGGGCACCTGCGTCTCCCAGTATGCCAACTGGCTGGTGGTGCTGCTCCTCGTCATCT
  A  G  T  C  V  S  Q  Y  A  N  W  L  V  V  L  L  L  V  I  F
     3130                3150                3170
TCCTGCTCGTGGCCAACATCCTGCTGGTCAACTTGCTCATTGCCATGTTCAGTTACACAT
  L  L  V  A  N  I  L  L  V  N  L  L  I  A  M  F  S  Y  T  F
     3190                3210                3230
TCGGCAAAGTACAGGGCAACAGCGATCTCTACTGGAAGGCGCAGCGTTACCGCCTCATCC
  G  K  V  Q  G  N  S  D  L  Y  W  K  A  Q  R  Y  R  L  I  R
     3250                3270                3290
GGGAATTCCACTCTCGGCCCGCGCTGGCCCCGCCCTTTATCGTCATCTCCCACTTGCGCC
  E  F  H  S  R  P  A  L  A  P  P  F  I  V  I  S  H  L  R  L
     3310                3330                3350
TCCTGCTCAGGCAATTGTGCAGGCGACCCCGGAGCCCCCAGCCGTCCTCCCCGGCCCTCG
  L  L  R  Q  L  C  R  R  P  R  S  P  Q  P  S  S  P  A  L  E
     3370                3390                3410
AGCATTTCCGGGTTTACCTTTCTAAGGAAGCCGAGCGGAAGCTGCTAACGTGGGAATCGG
  H  F  R  V  Y  L  S  K  E  A  E  R  K  L  L  T  W  E  S  V
     3430                3450                3470
TGCATAAGGAGAACTTTCTGCTGGCACGCGCTAGGGACAAGCGGGAGAGCGACTCCGAGC
  H  K  E  N  F  L  L  A  R  A  R  D  K  R  E  S  D  S  E  R
     3490                3510                3530
GTCTGAAGCGCACGTCCCAGAAGGTGGACTTGGCACTGAAACAGCTGGGACACATCCGCG
  L  K  R  T  S  Q  K  V  D  L  A  L  K  Q  L  G  H  I  R  E
```

Fig. 9 / continua ı 3

```
      3550                3570                3590
AGTACGAACAGCGCCTGAAAGTGCTGGAGCGGGAGGTCCAGCAGTGTAGCCGCGTCCTGG
  Y  E  Q  R  L  K  V  L  E  R  E  V  Q  Q  C  S  R  V  L  G
      3610                3630                3650
GGTGGGTGGCCGAGGCCCTGAGCCGCTCTGCCTTGCTGCCCCCAGGTGGGCCGCCACCCC
  W  V  A  E  A  L  S  R  S  A  L  L  P  P  G  G  P  P  P  P
      3670                3690                3710
CTGACCTGCCTGGGTCCAAAGACTGAGCCCTGCTGGCGGACTTCAAGGAGAAGCCCCCAC
  D  L  P  G  S  K  D  *
      3730                3750                3770
AGGGGATTTTGCTCCTAGAGTAAGGCTCATCTGGGCCTCGGCCCCCGCACCTGGTGGCCT
      3790                3810                3830
TGTCCTTGAGGTGAGCCCCATGTCCATCTGGGCCACTGTCAGGACCACCTTTGGGAGTGT
      3850                3870                3890
CATCCTTACAAACCACAGCATGCCCGGCTCCTCCCAGAACCAGTCCCAGCCTGGGAGGAT
      3910                3930                3950
CAAGGCCTGGATCCCGGGCCGTTATCCATCTGGAGGCTGCAGGGTCCTTGGGGTAACAGG
      3970                3990                4010
GACCACAGACCCCTCACCACTCACAGATTCCTCACACTGGGGAAATAAAGCCATTTCAGA
      4030
GGAAAAAAAAAAAAAAAAAAAA
```

```
MVVPEKEQSWIPKIFKKKTCTTFIVDSTDPGGTLCQCGRPRTAHPAVAMEDAFGAAVVTVWDSDAHTTEKPTDAYELDFTGAGRKH
SNFLRLSDRTDPAAVYSLVTRTWGFRAPNLVVSVLGGSGGPVLQTWLQDLLRRGLVRAAQSTGAWIVTGGLHTGIGRHVGVAVRDH
QMASTGGTKVVAMGVAPWGVVRNRDTLINPKGSFPARYRWRGDPEDGVQFPLDYNYSAFFLVDDGTHGCLGGENRFRLRLESYISQ
QKTGVGGTGIDIPVLLLLIDGDEKMLTRIENATQAHVPCLLVAGSRGLGNPGGTLEAHLAQDGDHKANQSTNQLLLPKDLSLQPVE
SIDRKTLQSYSERLAVAWNRVDIAQSELFRGDIQWRSFHLEASLMDALLNDRPEFVRLLISHGLSLGHFLTPMRLAQLYSAAPSNS
LIRNLLDQASHSAGTKAPALKGGAAELRPPDVGHVLRMLLGKMCAPRYPSGGAWDPHPGQGFGESMYLLSDKATSPLSLDAGLGQA
PWSDLLLWALLLNRAQMAMYFWEMGSNAVSSALGACLLLRVMARLEPDAEEAARRKDLAFKFEGMGVDLFGECYRSSEVRAARLLL
RRCPLWGDATCLQLAMQADARAFFAQDGVQSLLTQKWWGDMASTTPIWALVLAFFCPPLIYTRLITFRKSEEEPTREELEFDMDSV
INGEGPVGTADPAEKTPLGVPRQSGRPGCCGGRCGGRRCLRRWFHFWGVPVTIFMGNVVSYLLFLLLFSRVLLVDFQPAPPGSLEL
LLYFWAFTLLCEELRQGLSGGGGSLASGGPGPGHASLSQRLRLYLADSWNQCDLVALTCFLLGVGCRLTPGLYHLGRTVLCIDFMV
FTVRLLHIFTVNKQLGPKIVIVSKMMKDVFFFLFFLGVWLVAYGVATEGLLRPRDSDFPSILRRVFYRPYLQIFGQIPQEDMDVAL
MEHSNCSSEPGFWAHPPGAQAGTCVSQYANWLVVLLLVIFLLVANILLVNLLIAMFSYTFGKVQGNSDLYWKAQRYRLIREFHSRP
ALAPPFIVISHLRLLLRQLCRRPRSPQPSSPALEHFRVYLSKEAERKLLTWESVHKENFLLARARDKRESDSERLKRTSQKVDLAL
KQLGHIREYEQRLKVLEREVQQCSRVLGWVAEALSRSALLPPGGPPPPDLPGSKD
```

B.)

```
       10                  30                  50
ATCCAATGGCGGTCCTTCCATCTCGAAGCTTCCCTCATGGACGCCCTGCTGAATGACCGG
       70                  90                 110
CCTGAGTTCGTGCGCTTGCTCATTTCCCACGGCCTCAGCCTGGGCCACTTCCTGACCCCG
      130                 150                 170
ATGCGCCTGGCCCAACTCTACAGCGCGGCGCCCTCCAACTCGCTCATCCGCAACCTTTTG
      190                 210                 230
GACCAGGCGTCCCACAGCGCAGGCACCAAAGCCCCAGCCCTAAAAGGGGGAGCTGCGGAG
      250                 270                 290
CTCCGGCCCCCTGACGTGGGGCATGTGCTGAGGATGCTGCTGGGGAAGATGTGCGCGCCG
      310                 330                 350
AGATGTATCTGCTCTCGGACAAGGCCACCTCGCCGCTCTCGCTGGATGCTGGCCTCGGGC
  M   Y   L   L   S   D   K   A   T   S   P   L   S   L   D   A   G   L   G   Q
      370                 390                 410
AGGCCCCCTGGAGCGACCTGCTTCTTTGGGCACTGTTGCTGAACAGGGCACAGATGGCCA
  A   P   W   S   D   L   L   L   W   A   L   L   L   N   R   A   Q   M   A   M
      430                 450                 470
TGTACTTCTGGGAGATGGGTTCCAATGCAGTTTCCTCAGCTCTTGGGGCCTGTTTGCTGC
  Y   F   W   E   M   G   S   N   A   V   S   S   A   L   G   A   C   L   L   L
```

Fig. 9 / continue 1 4

```
      490                 510                 530
TCCGGGTGATGGCACGCCTGGAGCCTGACGCTGAGGAGGCAGCACGGAGGAAAGACCTGG
  R  V  M  A  R  L  E  P  D  A  E  E  A  A  R  R  K  D  L  A
      550                 570                 590
CGTTCAAGTTTGAGGGGATGGGCGTTGACCTCTTTGGCGAGTGCTATCGCAGCAGTGAGG
  F  K  F  E  G  M  G  V  D  L  F  G  E  C  Y  R  S  S  E  V
      610                 630                 650
TGAGGGCTGCCCGCCTCCTCCTCCGTCGCTGCCCGCTCTGGGGGGATGCCACTTGCCTCC
  R  A  A  R  L  L  L  R  R  C  P  L  W  G  D  A  T  C  L  Q
      670                 690                 710
AGCTGGCCATGCAAGCTGACGCCCGTGCCTTCTTTGCCCAGGATGGGGTACAGTCTCTGC
  L  A  M  Q  A  D  A  R  A  F  F  A  Q  D  G  V  Q  S  L  L
      730                 750                 770
TGACACAGAAGTGGTGGGGAGATATGGCCAGCACTACACCCATCTGGGCCCTGGTTCTCG
  T  Q  K  W  W  G  D  M  A  S  T  T  P  I  W  A  L  V  L  A
      790                 810                 830
CCTTCTTTTGCCCTCCACTCATCTACACCCGCCTCATCACCTTCAGGAAATCAGAAGAGG
  F  F  C  P  P  L  I  Y  T  R  L  I  T  F  R  K  S  E  E  E
      850                 870                 890
AGCCCACACGGGAGGAGCTAGAGTTTGACATGGATAGTGTCATTAATGGGGAAGGGCCTG
  P  T  R  E  E  L  E  F  D  M  D  S  V  I  N  G  E  G  P  V
      910                 930                 950
TCGGGACGGCGGACCCAGCCGAGAAGACGCCGCTGGGGGTCCCGCGCCAGTCGGGCCGTC
  G  T  A  D  P  A  E  K  T  P  L  G  V  P  R  Q  S  G  R  P
      970                 990                1010
CGGGTTGCTGCGGGGGCCGCTGCGGGGGGCGCCGGTGCCTACGCCGCTGGTTCCACTTCT
  G  C  C  G  G  R  C  G  G  R  R  C  L  R  R  W  F  H  F  W
     1030                1050                1070
GGGGCGTGCCGGTGACCATCTTCATGGGCAACGTGGTCAGCTACCTGCTGTTCCTGCTGC
  G  V  P  V  T  I  F  M  G  N  V  V  S  Y  L  L  F  L  L  L
     1090                1110                1130
TTTTCTCGCGGGTGCTGCTCGTGGATTTCCAGCCGGCGCCGCCCGGCTCCCTGGAGCTGC
  F  S  R  V  L  L  V  D  F  Q  P  A  P  P  G  S  L  E  L  L
     1150                1170                1190
TGCTCTATTTCTGGGCTTTCACGCTGCTGTGCGAGGAACTGCGCCAGGGCCTGAGCGGAG
  L  Y  F  W  A  F  T  L  L  C  E  E  L  R  Q  G  L  S  G  G
     1210                1230                1250
GCGGGGGCAGCCTCGCCAGCGGGGGCCCCGGGCCTGGCCATGCCTCACTGAGCCAGCGCC
  G  G  S  L  A  S  G  G  P  G  P  G  H  A  S  L  S  Q  R  L
     1270                1290                1310
TGCGCCTCTACCTCGCCGACAGCTGGAACCAGTGCGACCTAGTGGCTCTCACCTGCTTCC
  R  L  Y  L  A  D  S  W  N  Q  C  D  L  V  A  L  T  C  F  L
     1330                1350                1370
TCCTGGGCGTGGGCTGCCGGCTGACCCCGGGTTTGTACCACCTGGGCCGCACTGTCCTCT
  L  G  V  G  C  R  L  T  P  G  L  Y  H  L  G  R  T  V  L  C
     1390                1410                1430
GCATCGACTTCATGGTTTTCACGGTGCGGCTGCTTCACATCTTCACGGTCAACAAACAGC
  I  D  F  M  V  F  T  V  R  L  L  H  I  F  T  V  N  K  Q  L
     1450                1470                1490
TGGGGCCCAAGATCGTCATCGTGAGCAAGATGATGAAGGACGTGTTCTTCTTCCTCTTCT
  G  P  K  I  V  I  V  S  K  M  M  K  D  V  F  F  F  L  F  F
     1510                1530                1550
TCCTCGGCGTGTGGCTGGTAGCCTATGGCGTGGCCACGGAGGGGCTCCTGAGGCCACGGG
  L  G  V  W  L  V  A  Y  G  V  A  T  E  G  L  L  R  P  R  D
     1570                1590                1610
ACAGTGACTTCCCAAGTATCCTGCGCCGCGTCTTCTACCGTCCCTACCTGCAGATCTTCG
  S  D  F  P  S  I  L  R  R  V  F  Y  R  P  Y  L  Q  I  F  G
     1630                1650                1670
GGCAGATTCCCCAGGAGGACATGGACGTGGCCCTCATGGAGCACAGCAACTGCTCGTCGG
  Q  I  P  Q  E  D  M  D  V  A  L  M  E  H  S  N  C  S  S  E
     1690                1710                1730
AGCCCGGCTTCTGGGCACACCCTCCTGGGGCCCAGGCGGGCACCTGCGTCTCCCAGTATG
```

Fig. 9 / continuation 5

```
  P  G  F  W  A  H  P  P  G  A  Q  A  G  T  C  V  S  Q  Y  A
     1750                 1770                 1790
CCAACTGGCTGGTGGTGCTGCTCCTCGTCATCTTCCTGCTCGTGGCCAACATCCTGCTGG
  N  W  L  V  V  L  L  L  V  I  F  L  L  V  A  N  I  L  L  V
     1810                 1830                 1850
TCAACTTGCTCATTGCCATGTTCAGTTACACATTCGGCAAAGTACAGGGCAACAGCGATC
  N  L  L  I  A  M  F  S  Y  T  F  G  K  V  Q  G  N  S  D  L
     1870                 1890                 1910
TCTACTGGAAGGCGCAGCGTTACCGCCTCATCCGGGAATTCCACTCTCGGCCCGCGCTGG
  Y  W  K  A  Q  R  Y  R  L  I  R  E  F  H  S  R  P  A  L  A
     1930                 1950                 1970
CCCCGCCCTTTATCGTCATCTCCCACTTGCGCCTCCTGCTCAGGCAATTGTGCAGGCGAC
  P  P  F  I  V  I  S  H  L  R  L  L  L  R  Q  L  C  R  R  P
     1990                 2010                 2030
CCCGGAGCCCCCAGCCGTCCTCCCCGGCCCTCGAGCATTTCCGGGTTTACCTTTCTAAGG
  R  S  P  Q  P  S  S  P  A  L  E  H  F  R  V  Y  L  S  K  E
     2050                 2070                 2090
AAGCCGAGCGGAAGCTGCTAACGTGGGAATCGGTGCATAAGGAGAACTTTCTGCTGGCAC
  A  E  R  K  L  L  T  W  E  S  V  H  K  E  N  F  L  L  A  R
     2110                 2130                 2150
GCGCTAGGGACAAGCGGGAGAGCGACTCCGAGCGTCTGAAGCGCACGTCCCAGAAGGTGG
  A  R  D  K  R  E  S  D  S  E  R  L  K  R  T  S  Q  K  V  D
     2170                 2190                 2210
ACTTGGCACTGAAACAGCTGGGACACATCCGCGAGTACGAACAGCGCCTGAAAGTGCTGG
  L  A  L  K  Q  L  G  H  I  R  E  Y  E  Q  R  L  K  V  L  E
     2230                 2250                 2270
AGCGGGAGGTCCAGCAGTGTAGCCGCGTCCTGGGGTGGGTGGCCGAGGCCCTGAGCCGCT
  R  E  V  Q  Q  C  S  R  V  L  G  W  V  A  E  A  L  S  R  S
     2290                 2310                 2330
CTGCCTTGCTGCCCCCAGGTGGGCCGCCACCCCCTGACCTGCCTGGGTCCAAAGACTGAG
  A  L  L  P  P  G  G  P  P  P  P  D  L  P  G  S  K  D  *
     2350                 2370                 2390
CCCTGCTGGCGGACTTCAAGGAGAAGCCCCCACAGGGGATTTTGCTCCTAGAGTAAGGCT
     2410                 2430                 2450
CATCTGGGCCTCGGCCCCCGCACCTGGTGGCCTTGTCCTTGAGGTGAGCCCCATGTCCAT
     2470                 2490                 2510
CTGGGCCACTGTCAGGACCACCTTTGGGAGTGTCATCCTTACAAACCACAGCATGCCCGG
     2530                 2550                 2570
CTCCTCCCAGAACCAGTCCCAGCCTGGGAGGATCAAGGCCTGGATCCCGGGCCGTTATCC
     2590                 2610                 2630
ATCTGGAGGCTGCAGGGTCCTTGGGGTAACAGGGACCACAGACCCCTCACCACTCACAGA
     2650                 2670                 2690
TTCCTCACACTGGGGAAATAAAGCCATTTCAGAGGAAAAAAAAAAAAAAAAAAAA
```

```
MYLLSDKATSPLSLDAGLGQAPWSDLLLWALLLNRAQMAMYFWEMGSNAVSSALGACLLLRVMARLEPDAEEAARRKDLAFKFEGM
GVDLFGECYRSSEVRAARLLLRRCPLWGDATCLQLAMQADARAFFAQDGVQSLLTQKWWGDMASTTPIWALVLAFFCPPLIYTRLI
TFRKSEEEPTREELEFDMDSVINGEGPVGTADPAEKTPLGVPRQSGRPGCCGGRCGGRRCLRRWFHFWGVPVTIFMGNVVSYLLFL
LLFSRVLLVDFQPAPPGSLELLLYFWAFTLLCEELRQGLSGGGGSLASGGPGPGHASLSQRLRLYLADSWNQCDLVALTCFLLGVG
CRLTPGLYHLGRTVLCIDFMVFTVRLLHIFTVNKQLGPKIVIVSKMMKDVFFFLFFLGVWLVAYGVATEGLLRPRDSDFPSILRRV
FYRPYLQIFGQIPQEDMDVALMEHSNCSSEPGFWAHPPGAQAGTCVSQYANWLVVLLLVIFLLVANILLVNLLIAMFSYTFGKVQG
NSDLYWKAQRYRLIREFHSRPALAPPFIVISHLRLLLRQLCRRPRSPQPSSPALEHFRVYLSKEAERKLLTWESVHKENFLLARAR
DKRESDSERLKRTSQKVDLALKQLGHIREYEQRLKVLEREVQQCSRVLGWVAEALSRSALLPPGGPPPPDLPGSKD
```

A)

```
        10                  30                  50
ATTAAAGTTTATAAAACAGTGGCTGGATGGTTGGAGGATGCAGGTGGACAGAAGACGTGG
                          M  V  G  G  C  R  W  T  E  D  V  E
        70                  90                 110
AGCCTGCAGAAGTAAAGGAAAAGATGTCCTTTCGGGCAGCCAGGCTCAGCATGAGGAACA
  P  A  E  V  K  E  K  M  S  F  R  A  A  R  L  S  M  R  N  R
       130                 150                 170
GAAGGAATGACACTCTGGACAGCACCCGGACCCTGTACTCCAGCGCGTCTCGGAGCACAG
  R  N  D  T  L  D  S  T  R  T  L  Y  S  S  A  S  R  S  T  D
       190                 210                 230
ACTTGTCTTACAGTGAAAGCGCCAGCTTCTACGCTGCCTTCAGGACACAGACGTGCCCAA
  L  S  Y  S  E  S  A  S  F  Y  A  A  F  R  T  Q  T  C  P  I
       250                 270                 290
TCATGGCTTCTTGGGACTTGGTGAATTTTATTCAAGCAAATTTTAAGAAACGAGAATGTG
  M  A  S  W  D  L  V  N  F  I  Q  A  N  F  K  K  R  E  C  V
       310                 330                 350
TCTTCTTTACCAAAGATTCCAAGGCCACGGAGAATGTGTGCAAGTGTGGCTATGCCCAGA
  F  F  T  K  D  S  K  A  T  E  N  V  C  K  C  G  Y  A  Q  S
       370                 390                 410
GCCAGCACATGGAAGGCACCCAGATCAACCAAAGTGAGAAATGGAACTACAAGAAACACA
  Q  H  M  E  G  T  Q  I  N  Q  S  E  K  W  N  Y  K  K  H  T
       430                 450                 470
CCAAGGAATTTCCTACCGACGCCTTTGGGGATATTCAGTTTGAGACACTGGGGAAGAAAG
  K  E  F  P  T  D  A  F  G  D  I  Q  F  E  T  L  G  K  K  G
       490                 510                 530
GGAAGTATATACGTCTGTCCTGCGACACGGACGCGGAAATCCTTTACGAGCTGCTGACCC
  K  Y  I  R  L  S  C  D  T  D  A  E  I  L  Y  E  L  L  T  Q
       550                 570                 590
AGCACTGGCACCTGAAAACACCCAACCTGGTCATTTCTGTGACCGGGGGGCGCCAAGAACT
  H  W  H  L  K  T  P  N  L  V  I  S  V  T  G  G  A  K  N  F
       610                 630                 650
TCGCCCTGAAGCCGCGCATGCGCAAGATCTTCAGCCGGCTCATCTACATCGCGCAGTCCA
  A  L  K  P  R  M  R  K  I  F  S  R  L  I  Y  I  A  Q  S  K
       670                 690                 710
AAGGTGCTTGGATTCTCACGGGAGGCACCCATTATGGCCTGATGAAGTACATCGGGGAGG
  G  A  W  I  L  T  G  G  T  H  Y  G  L  M  K  Y  I  G  E  V
       730                 750                 770
TGGTGAGAGATAACACCATCAGCAGGAGTTCAGAGGAGAATATTGTGGCCATTGGCATAG
  V  R  D  N  T  I  S  R  S  S  E  E  N  I  V  A  I  G  I  A
       790                 810                 830
CAGCTTGGGGCATGGTCTCCAACCGGGACACCCTCATCAGGAATTGCGATGCTGAGGGCT
  A  W  G  M  V  S  N  R  D  T  L  I  R  N  C  D  A  E  G  Y
       850                 870                 890
ATTTTTTAGCCCAGTACCTTATGGATGACTTCACAAGAGATCCACTGTATATCCTGGACA
  F  L  A  Q  Y  L  M  D  D  F  T  R  D  P  L  Y  I  L  D  N
       910                 930                 950
ACAACCACACACATTTGCTGCTCGTGGACAATGGCTGTCATGGACATCCCACTGTCGAAG
  N  H  T  H  L  L  L  V  D  N  G  C  H  G  H  P  T  V  E  A
       970                1990                1010
CAAAGCTCCGGAATCAGCTAGAGAAGTATATCTCTGAGCGCACTATTCAAGATTCCAACT
  K  L  R  N  Q  L  E  K  Y  I  S  E  R  T  I  Q  D  S  N  Y
      1030                1050                1070
ATGGTGGCAAGATCCCCATTGTGTGTTTTGCCCAAGGAGGTGGAAAAGAGACTTTGAAAG
  G  G  K  I  P  I  V  C  F  A  Q  G  G  G  K  E  T  L  K  A
      1090                1110                1130
CCATCAATACCTCCATCAAAAATAAAATTCCTTGTGTGGTGGTGGAAGGCTCGGGCCAGA
  I  N  T  S  I  K  N  K  I  P  C  V  V  V  E  G  S  G  Q  I
      1150                1170                1190
TCGCTGATGTGATCGCTAGCCTGGTGGAGGTGGAGGATGCCCTGACATCTTCTGCCGTCA
  A  D  V  I  A  S  L  V  E  V  E  D  A  L  T  S  S  A  V  K
      1210                1230                1250
```

Fig. 10 / continuation 1

```
AGGAGAAGCTGGTGCGCTTTTTACCCCGCACGGTGTCCCGGCTGCCTGAGGAGGAGACTG
  E  K  L  V  R  F  L  P  R  T  V  S  R  L  P  E  E  E  T  E
     1270              1290              1310
AGAGTTGGATCAAATGGCTCAAAGAAATTCTCGAATGTTCTCACCTATTAACAGTTATTA
  S  W  I  K  W  L  K  E  I  L  E  C  S  H  L  L  T  V  I  K
     1330              1350              1370
AAATGGAAGAAGCTGGGGATGAAATTGTGAGCAATGCCATCTCCTACGCTCTATACAAAG
  M  E  E  A  G  D  E  I  V  S  N  A  I  S  Y  A  L  Y  K  A
     1390              1410              1430
CCTTCAGCACCAGTGAGCAAGACAAGGATAACTGGAATGGGCAGCTGAAGCTTCTGCTGG
  F  S  T  S  E  Q  D  K  D  N  W  N  G  Q  L  K  L  L  L  E
     1450              1470              1490
AGTGGAACCAGCTGGACTTAGCCAATGATGAGATTTTCACCAATGACCGCCGATGGGAGA
  W  N  Q  L  D  L  A  N  D  E  I  F  T  N  D  R  R  W  E  K
     1510              1530              1550
AGAGCAAACCGAGGCTCAGAGACACAATAATCCAGGTCACATGGCTGGAAAATGGTAGAA
  S  K  P  R  L  R  D  T  I  I  Q  V  T  W  L  E  N  G  R  I
     1570              1590              1610
TCAAGGTTGAGAGCAAAGATGTGACTGACGGCAAAGCCTCTTCTCATATGCTGGTGGTTC
  K  V  E  S  K  D  V  T  D  G  K  A  S  S  H  M  L  V  V  L
     1630              1650              1670
TCAAGTCTGCTGACCTTCAAGAAGTCATGTTTACGGCTCTCATAAAGGACAGACCCAAGT
  K  S  A  D  L  Q  E  V  M  F  T  A  L  I  K  D  R  P  K  F
     1690              1710              1730
TTGTCCGCCTCTTTCTGGAGAATGGCTTGAACCTACGGAAGTTTCTCACCCATGATGTCC
  V  R  L  F  L  E  N  G  L  N  L  R  K  F  L  T  H  D  V  L
     1750              1770              1790
TCACTGAACTCTTCTCCAACCACTTCAGCACGCTTGTGTACCGGAATCTGCAGATCGCCA
  T  E  L  F  S  N  H  F  S  T  L  V  Y  R  N  L  Q  I  A  K
     1810              1830              1850
AGAATTCCTATAATGATGCCCTCCTCACGTTTGTCTGGAAACTGGTTGCGAACTTCCGAA
  N  S  Y  N  D  A  L  L  T  F  V  W  K  L  V  A  N  F  R  R
     1870              1890              1910
GAGGCTTCCGGAAGGAAGACAGAAATGGCCGGGACGAGATGGACATAGAACTCCACGACG
  G  F  R  K  E  D  R  N  G  R  D  E  M  D  I  E  L  H  D  V
     1930              1950              1970
TGTCTCCTATTACTCGGCACCCCCTGCAAGCTCTCTTCATCTGGGCCATTCTTCAGAATA
  S  P  I  T  R  H  P  L  Q  A  L  F  I  W  A  I  L  Q  N  K
     1990              2010              2030
AGAAGGAACTCTCCAAAGTCATTTGGGAGCAGACCAGGGGCTGCACTCTGGCAGCCCTGG
  K  E  L  S  K  V  I  W  E  Q  T  R  G  C  T  L  A  A  L  G
     2050              2070              2090
GAGCCAGCAAGCTTCTGAAGACTCTGGCCAAAGTGAAGAACGACATCAATGCTGCTGGGG
  A  S  K  L  L  K  T  L  A  K  V  K  N  D  I  N  A  A  G  E
     2110              2130              2150
AGTCCGAGGAGCTGGCTAATGAGTACGAGACCCGGGCTGTTGGTGAGTCCACAGTGTGGA
  S  E  E  L  A  N  E  Y  E  T  R  A  V  G  E  S  T  V  W  N
     2170              2190              2210
ATGCTGTGGTGGGCGCGGATCTGCCATGTGGCACAGACATTGCCAGCGGCACTCATAGAC
  A  V  V  G  A  D  L  P  C  G  T  D  I  A  S  G  T  H  R  P
     2230              2250              2270
CAGATGGTGGAGAGCTGTTCACTGAGTGTTACAGCAGCGATGAAGACTTGGCAGAACAGC
  D  G  G  E  L  F  T  E  C  Y  S  S  D  E  D  L  A  E  Q  L
     2290              2310              2330
TGCTGGTCTATTCCTGTGAAGCTTGGGGTGGAAGCAACTGTCTGGAGCTGGCGGTGGAGG
  L  V  Y  S  C  E  A  W  G  G  S  N  C  L  E  L  A  V  E  A
     2350              2370              2390
CCACAGACCAGCATTTCATCGCCCAGCCTGGGGTCCAGAATTTTCTTTCTAAGCAATGGT
  T  D  Q  H  F  I  A  Q  P  G  V  Q  N  F  L  S  K  Q  W  Y
     2410              2430              2450
ATGGAGAGATTTCCCGAGACACCAAGAACTGGAAGATTATCCTGTGTCTGTTTATTATAC
  G  E  I  S  R  D  T  K  N  W  K  I  I  L  C  L  F  I  I  P
```

Fig. 10 / continuation 2

```
     2470                2490                2510
CCTTGGTGGGCTGTGGCTTTGTATCATTTAGGAAGAAACCTGTCGACAAGCACAAGAAGC
  L  V  G  C  G  F  V  S  F  R  K  K  P  V  D  K  H  K  K  L
     2530                2550                2570
TGCTTTGGTACTATGTGGCGTTCTTCACCTCCCCCTTCGTGGTCTTCTCCTGGAATGTGG
  L  W  Y  Y  V  A  F  F  T  S  P  F  V  V  F  S  W  N  V  V
     2590                2610                2630
TCTTCTACATCGCCTTCCTCCTGCTGTTTGCCTACGTGCTGCTCATGGATTTCCATTCGG
  F  Y  I  A  F  L  L  L  F  A  Y  V  L  L  M  D  F  H  S  V
     2650                2670                2690
TGCCACACCCCCCCGAGCTGGTCCTGTACTCGCTGGTCTTTGTCCTCTTCTGTGATGAAG
  P  H  P  P  E  L  V  L  Y  S  L  V  F  V  L  F  C  D  E  V
     2710                2730                2750
TGAGACAGGGCCGGCCGGCTGCTCCCAGTGCGGGGCCCGCCAAGCCCACGCCCACCCGGA
  R  Q  G  R  P  A  A  P  S  A  G  P  A  K  P  T  P  T  R  N
     2770                2790                2810
ACTCCATCTGGCCCGCAAGCTCCACACGCAGCCCCGGTTCCCGCTCACGCCACTCCTTCC
  S  I  W  P  A  S  S  T  R  S  P  G  S  R  S  R  H  S  F  H
     2830                2850                2870
ACACTTCCCTGCAAGCTGAGGGTGCCAGCTCTGGCCTTGGCCAGCCCAGAAAGGGGTGGA
  T  S  L  Q  A  E  G  A  S  S  G  L  G  Q  P  R  K  G  W  T
     2890                2910                2930
CATTTAAAAATCTGGAAATGGTTGATATTTCCAAGCTGCTGATGTCCCTCTCTGTCCCTT
  F  K  N  L  E  M  V  D  I  S  K  L  L  M  S  L  S  V  P  F
     2950                2970                2990
TCTGTACGCAGTGGTACGTAAATGGGGTGAATTATTTTACTGACCTGTGGAATGTGATGG
  C  T  Q  W  Y  V  N  G  V  N  Y  F  T  D  L  W  N  V  M  D
     3010                3030                3050
ACACGCTGGGGCTTTTTTACTTCATAGCAGGAATTGTATTTCGGCAAGGGATCCTTAGGC
  T  L  G  L  F  Y  F  I  A  G  I  V  F  R  Q  G  I  L  R  Q
     3070                3090                3110
AGAATGAGCAGCGCTGGAGGTGGATATTCCGTTCGGTCATCTACGAGCCCTACCTGGCCA
  N  E  Q  R  W  R  W  I  F  R  S  V  I  Y  E  P  Y  L  A  M
     3130                3150                3170
TGTTCGGCCAGGTGCCCAGTGACGTGGATGGTACCACGTATGACTTTGCCCACTGCACCT
  F  G  Q  V  P  S  D  V  D  G  T  T  Y  D  F  A  H  C  T  F
     3190                3210                3230
TCACTGGGAATGAGTCCAAGCCACTGTGTGTGGAGCTGGATGAGCACAACCTGCCCCGGT
  T  G  N  E  S  K  P  L  C  V  E  L  D  E  H  N  L  P  R  F
     3250                3270                3290
TCCCCGAGTGGATCACCATCCCCCTGGTGTGCATCTACATGTTATCCACCAACATCCTGC
  P  E  W  I  T  I  P  L  V  C  I  Y  M  L  S  T  N  I  L  L
     3310                3330                3350
TGGTCAACCTGCTGGTCGCCATGTTTGGCTACACGGTGGGCACCGTCCAGGAGAACAATG
  V  N  L  L  V  A  M  F  G  Y  T  V  G  T  V  Q  E  N  N  D
     3370                3390                3410
ACCAGGTCTGGAAGTTCCAGAGGTACTTCCTGGTGCAGGAGTACTGCAGCCGCCTCAATA
  Q  V  W  K  F  Q  R  Y  F  L  V  Q  E  Y  C  S  R  L  N  I
     3430                3450                3470
TCCCCTTCCCCTTCATCGTCTTCGCTTACTTCTACATGGTGGTGAAGAAGTGCTTCAAGT
  P  F  P  F  I  V  F  A  Y  F  Y  M  V  V  K  K  C  F  K  C
     3490                3510                3530
GTTGCTGCAAGGAGAAAAACATGGAGTCTTCTGTCTGCTGTGAGTGGTTTATCCATGTGT
  C  C  K  E  K  N  M  E  S  S  V  C  C  E  W  F  I  H  V  Y
     3550                3570                3590
ACTTGGGATCAGAAGCAGCGATTAATTTCAGGGAAGGATGCCTGCATCCAGTGATTGGAA
  L  G  S  E  A  A  I  N  F  R  E  G  C  L  H  P  V  I  G  S
     3610                3630                3650
GCTGGACCCCAGGCTGGCTGGTCTGGACATCCACACGCATTCTCACATGCAGTGCCGGCT
  W  T  P  G  W  L  V  W  T  S  T  R  I  L  T  C  S  A  G  W
     3670                3690                3710
GGCCAGCAGCAGGGAGTCTCAGTGTCACCACACATAGCAGCTGGGTTCCTGCAAAAAGCA
```

Fig. 10 / continuation 3

```
  P  A  A  G  S  L  S  V  T  T  H  S  S  W  V  P  A  K  S  S
     3730                 3750                 3770
GCAAGTCACAGGCCCACCCAGACAGAACGGGTAGAGAATGTGACTCTGCTTCTGGGTGGG
  K  S  Q  A  H  P  D  R  T  G  R  E  C  D  S  A  S  G  W  E
     3790                 3810                 3830
AAGGACAGCCTGCCCGGTGGGTGGAAGAATCCGTGGCCCTGTTTGGCCATCGTGGCCCTG
  G  Q  P  A  R  W  V  E  E  S  V  A  L  F  G  H  R  G  P  V
     3850                 3870                 3890
TTTGGCCACCTACCACTCTAGGCATCACTGAGCTGAATGCGCCGGTCCTCTGA
  W  P  P  T  T  L  G  I  T  E  L  N  A  P  V  L  *
```

```
MVGGCRWTEDVEPAEVKEKMSFRAARLSMRNRRNDTLDSTRTLYSSASRSTDLSYSESASFYAAFRTQTCPIMASWDLVNFIQANF
KKRECVFFTKDSKATENVCKCGYAQSQHMEGTQINQSEKWNYKKHTKEFPTDAFGDIQFETLGKKGKYIRLSCDTDAEILYELLTQ
HWHLKTPNLVISVTGGAKNFALKPRMRKIFSRLIYIAQSKGAWILTGGTHYGLMKYIGEVVRDNTISRSSEENIVAIGIAAWGMVS
NRDTLIRNCDAEGYFLAQYLMDDFTRDPLYILDNNHTHLLLVDNGCHGHPTVEAKLRNQLEKYISERTIQDSNYGGKIPIVCFAQG
GGKETLKAINTSIKNKIPCVVVEGSGQIADVIASLVEVEDALTSSAVKEKLVRFLPRTVSRLPEEETESWIKWLKEILECSHLLTV
IKMEEAGDEIVSNAISYALYKAFSTSEQDKDNWNGQLKLLLEWNQLDLANDEIFTNDRRWEKSKPRLRDTIIQVTWLENGRIKVES
KDVTDGKASSHMLVVLKSADLQEVMFTALIKDRPKFVRLFLENGLNLRKFLTHDVLTELFSNHFSTLVYRNLQIAKNSYNDALLTF
VWKLVANFRRGFRKEDRNGRDEMDIELHDVSPITRHPLQALFIWAILQNKKELSKVIWEQTRGCTLAALGASKLLKTLAKVKNDIN
AAGESEELANEYETRAVGESTVWNAVVGADLPCGTDIASGTHRPDGGELFTECYSSDEDLAEQLLVYSCEAWGGSNCLELAVEATD
QHFIAQPGVQNFLSKQWYGEISRDTKNWKIILCLFIIPLVGCGFVSFRKKPVDKHKKLLWYYVAFFTSPFVVFSWNVVFYIAFLLL
FAYVLLMDFHSVPHPPELVLYSLVFVLFCDEVRQGRPAAPSAGPAKPTPTRNSIWPASSTRSPGSRSRHSFHTSLQAEGASSGLGQ
PRKGWTFKNLEMVDISKLLMSLSVPFCTQWYVNGVNYFTDLWNVMDTLGLFYFIAGIVFRQGILRQNEQRWRWIFRSVIYEPYLAM
FGQVPSDVDGTTYDFAHCTFTGNESKPLCVELDEHNLPRFPEWITIPLVCIYMLSTNILLVNLLVAMFGYTVGTVQENNDQVWKFQ
RYFLVQEYCSRLNIPFPFIVFAYFYMVVKKCFKCCCKEKNMESSVCCEWFIHVYLGSEAAINFREGCLHPVIGSWTPGWLVWTSTR
ILTCSAGWPAAGSLSVTTHSSWVPAKSSKSQAHPDRTGRECDSASGWEGQPARWVEESVALFGHRGPVWPPTTLGITELNAPVL
```

B.

```
                                                      Q  L
     2290                 2310                 2330
TGCTGGTCTATTCCTGTGAAGCTTGGGGTGGAAGCAACTGTCTGGAGCTGGCGGTGGAGG
  L  V  Y  S  C  E  A  W  G  G  S  N  C  L  E  L  A  V  E  A
     2350                 2370                 2390
CCACAGACCAGCATTTCATCGCCCAGCCTGGGGTCCAGAATTTTCTTTCTAAGCAATGGT
  T  D  Q  H  F  I  A  Q  P  G  V  Q  N  F  L  S  K  Q  W  Y
     2410                 2430                 2450
ATGGAGAGATTTCCCGAGACACCAAGAACTGGAAGATTATCCTGTGTCTGTTTATTATAC
  G  E  I  S  R  D  T  K  N  W  K  I  I  L  C  L  F  I  I  P
     2470                 2490                 2510
CCTTGGTGGGCTGTGGCTTTGTATCATTTAGGAAGAAACCTGTCGACAAGCACAAGAAGC
  L  V  G  C  G  F  V  S  F  R  K  K  P  V  D  K
```

Figure 11:

a.) Trp10b cDNA and derived amino acid sequence

```
        10                  30                  50
ATGAAATCCTTCCTTCCTGTCCACACCATCGTGCTTATCAGGGAGAATGTGTGCAAGTGT
M  K  S  F  L  P  V  H  T  I  V  L  I  R  E  N  V  C  K  C
        70                  90                  110
GGCTATGCCCAGAGCCAGCACATGGAAGGCACCCAGATCAACCAAAGTGAGAAATGGAAC
G  Y  A  Q  S  Q  H  M  E  G  T  Q  I  N  Q  S  E  K  W  N
       130                 150                 170
TACAAGAAACACACCAAGGAATTTCCTACCGACGCCTTTGGGGATATTCAGTTTGAGACA
Y  K  K  H  T  K  E  F  P  T  D  A  F  G  D  I  Q  F  E  T
       190                 210                 230
CTGGGGAAGAAAGGGAAGTATATACGTCTGTCCTGCGACACGGACGCGGAAATCCTTTAC
L  G  K  K  G  K  Y  I  R  L  S  C  D  T  D  A  E  I  L  Y
       250                 270                 290
GAGCTGCTGACCCAGCACTGGCACCTGAAAACACCCAACCTGGTCATTTCTGTGACCGGG
E  L  L  T  Q  H  W  H  L  K  T  P  N  L  V  I  S  V  T  G
       310                 330                 350
GGCGCCAAGAACTTCGCCCTGAAGCCGCGCATGCGCAAGATCTTCAGCCGGCTCATCTAC
G  A  K  N  F  A  L  K  P  R  M  R  K  I  F  S  R  L  I  Y
       370                 390                 410
ATCGCGCAGTCCAAAGGTGCTTGGATTCTCACGGGAGGCACCCATTATGGCCTGATGAAG
I  A  Q  S  K  G  A  W  I  L  T  G  G  T  H  Y  G  L  M  K
       430                 450                 470
TACATCGGGGAGGTGGTGAGAGATAACACCATCAGCAGGAGTTCAGAGGAGAATATTGTG
Y  I  G  E  V  V  R  D  N  T  I  S  R  S  S  E  E  N  I  V
       490                 510                 530
GCCATTGGCATAGCAGCTTGGGGCATGGTCTCCAACCGGGACACCCTCATCAGGAATTGC
A  I  G  I  A  A  W  G  M  V  S  N  R  D  T  L  I  R  N  C
       550                 570                 590
GATGCTGAGGGCTATTTTTTAGCCCAGTACCTTATGGATGACTTCACAAGAGATCCACTG
D  A  E  G  Y  F  L  A  Q  Y  L  M  D  D  F  T  R  D  P  L
       610                 630                 650
TATATCCTGGACAACAACCACACACATTTGCTGCTCGTGGACAATGGCTGTCATGGACAT
Y  I  L  D  N  N  H  T  H  L  L  L  V  D  N  G  C  H  G  H
       670                 690                 710
CCCACTGTCGAAGCAAAGCTCCGGAATCAGCTAGAGAAGTATATCTCTGAGCGCACTATT
P  T  V  E  A  K  L  R  N  Q  L  E  K  Y  I  S  E  R  T  I
       730                 750                 770
CAAGATTCCAACTATGGTGGCAAGATCCCCATTGTGTGTTTTGCCCAAGGAGGTGGAAAA
Q  D  S  N  Y  G  G  K  I  P  I  V  C  F  A  Q  G  G  G  K
       790                 810                 830
GAGACTTTGAAAGCCATCAATACCTCCATCAAAAATAAAATTCCTTGTGTGGTGGTGGAA
E  T  L  K  A  I  N  T  S  I  K  N  K  I  P  C  V  V  V  E
       850                 870                 890
GGCTCGGGCCAGATCGCTGATGTGATCGCTAGCCTGGTGGAGGTGGAGGATGCCCTGACA
G  S  G  Q  I  A  D  V  I  A  S  L  V  E  V  E  D  A  L  T
       910                 930                 950
TCTTCTGCCGTCAAGGAGAAGCTGGTGCGCTTTTTACCCCGCACGGTGTCCCGGCTGCCT
S  S  A  V  K  E  K  L  V  R  F  L  P  R  T  V  S  R  L  P
       970                 990                1010
GAGGAGGAGACTGAGAGTTGGATCAAATGGCTCAAAGAAATTCTCGAATGTTCTCACCTA
E  E  E  T  E  S  W  I  K  W  L  K  E  I  L  E  C  S  H  L
      1030                1050                1070
TTAACAGTTATTAAAATGGAAGAAGCTGGGGATGAAATTGTGAGCAATGCCATCTCCTAC
L  T  V  I  K  M  E  E  A  G  D  E  I  V  S  N  A  I  S  Y
      1090                1110                1130
GCTCTATACAAAGCCTTCAGCACCAGTGAGCAAGACAAGGATAACTGGAATGGGCAGCTG
A  L  Y  K  A  F  S  T  S  E  Q  D  K  D  N  W  N  G  Q  L
```

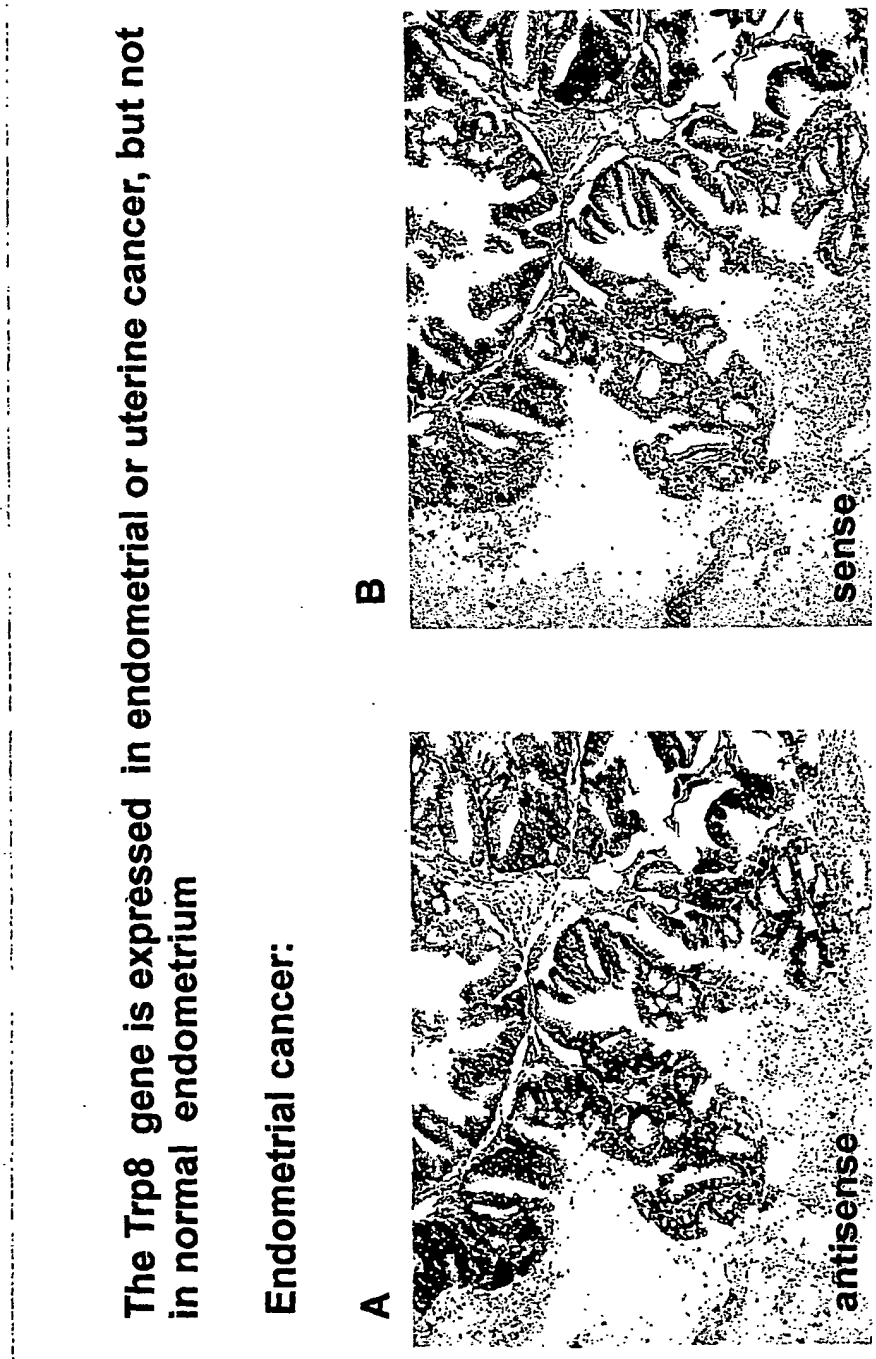
Fig. 11 (Continuation)

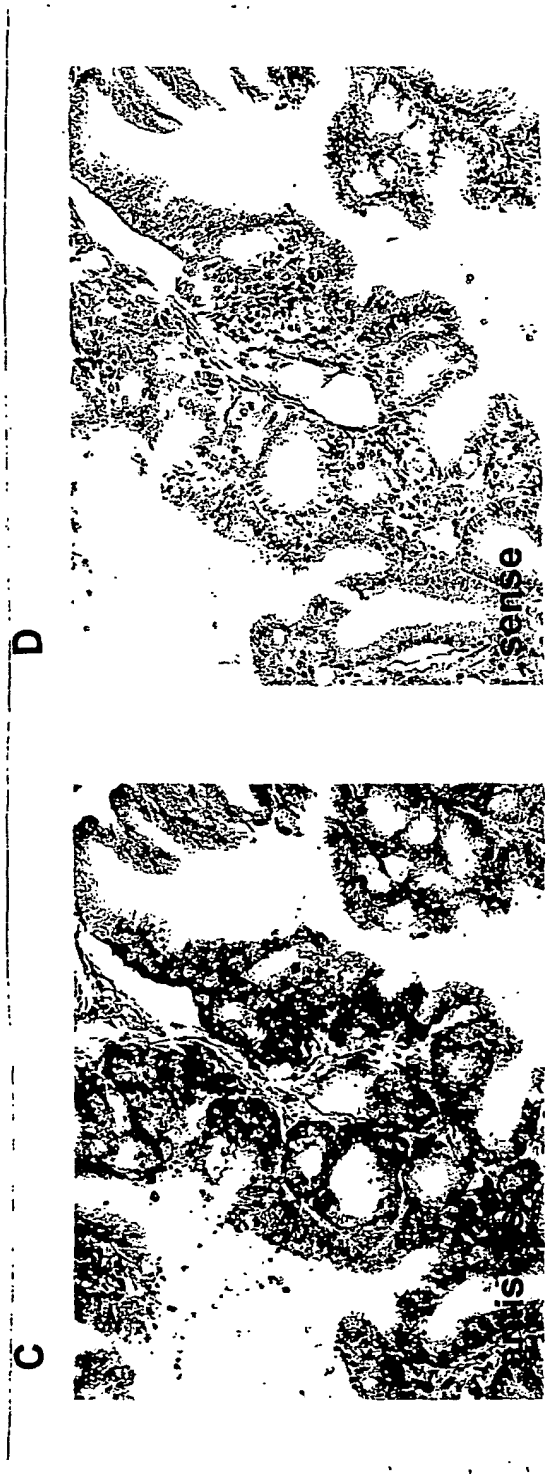
```
      2410                2430                2450
AGAAACTTAGGACCCAAGATTATAATGCTGCAGAGGATGCTGATCGATGTGTTCTTCTTC
R  N  L  G  P  K  I  I  M  L  Q  R  M  L  I  D  V  F  F  F
      2470                2490                2510
CTGTTCCTCTTTGCGGTGTGGATGGTGGCCTTTGGCGTGGCCAGGCAAGGGATCCTTAGG
L  F  L  F  A  V  W  M  V  A  F  G  V  A  R  Q  G  I  L  R
      2530                2550                2570
CAGAATGAGCAGCGCTGGAGGTGGATATTCCGTTCGGTCATCTACGAGCCCTACCTGGCC
Q  N  E  Q  R  W  R  W  I  F  R  S  V  I  Y  E  P  Y  L  A
      2590                2610                2630
ATGTTCGGCCAGGTGCCCAGTGACGTGGATGGTACCACGTATGACTTTGCCCACTGCACC
M  F  G  Q  V  P  S  D  V  D  G  T  T  Y  D  F  A  H  C  T
      2650                2670                2690
TTCACTGGGAATGAGTCCAAGCCACTGTGTGTGGAGCTGGATGAGCACAACCTGCCCCGG
F  T  G  N  E  S  K  P  L  C  V  E  L  D  E  H  N  L  P  R
      2710                2730                2750
TTCCCCGAGTGGATCACCATCCCCCTGGTGTGCATCTACATGTTATCCACCAACATCCTG
F  P  E  W  I  T  I  P  L  V  C  I  Y  M  L  S  T  N  I  L
      2770                2790                2810
CTGGTCAACCTGCTGGTCGCCATGTTTGGCTACACGGTGGGCACCGTCCAGGAGAACAAT
L  V  N  L  L  V  A  M  F  G  Y  T  V  G  T  V  Q  E  N  N
      2830                2850                2870
GACCAGGTCTGGAAGTTCCAGAGGTACTTCCTGGTGCAGGAGTACTGCAGCCGCCTCAAT
D  Q  V  W  K  F  Q  R  Y  F  L  V  Q  E  Y  C  S  R  L  N
      2890                2910                2930
ATCCCCTTCCCCTTCATCGTCTTCGCTTACTTCTACATGGTGGTGAAGAAGTGCTTCAAG
I  P  F  P  F  I  V  F  A  Y  F  Y  M  V  V  K  K  C  F  K
      2950                2970                2990
TGTTGCTGCAAGGAGAAAAACATGGAGTCTTCTGTCTGCTGTTTCAAAAATGAAGACAAT
C  C  C  K  E  K  N  M  E  S  S  V  C  C  F  K  N  E  D  N
      3010                3030                3050
GAGACTCTGGCATGGGAGGGTGTCATGAAGGAAAACTACCTTGTCAAGATCAACACAAAA
E  T  L  A  W  E  G  V  M  K  E  N  Y  L  V  K  I  N  T  K
      3070                3090                3110
GCCAACGACACCTCAGAGGAAATGAGGCATCGATTTAGACAACTGGATACAAAGCTTAAT
A  N  D  T  S  E  E  M  R  H  R  F  R  Q  L  D  T  K  L  N
      3130                3150
GATCTCAAGGGTCTACTGAAAGAGATTGCTAATAAAATCAAATAG
D  L  K  G  L  L  K  E  I  A  N  K  I  K  *
```

b.) Trp10 protein:

```
MKSFLPVHTIVLIRENVCKCGYAQSQHMEGTQINQSEKWNYKKHTKEFPTDAFGDIQFETLGKKGKYIRLSCDTDAEILY
ELLTQHWHLKTPNLVISVTGGAKNFALKPRMRKIFSRLIYIAQSKGAWILTGGTHYGLMKYIGEVVRDNTISRSSEENIV
AIGIAAWGMVSNRDTLIRNCDAEGYFLAQYLMDDFTRDPLYILDNNHTHLLLVDNGCHGHPTVEAKLRNQLEKYISERTI
QDSNYGGKIPIVCFAQGGGKETLKAINTSIKNKIPCVVVEGSGQIADVIASLVEVEDALTSSAVKEKLVRFLPRTVSRLP
EEETESWIKWLKEILECSHLLTVIKMEEAGDEIVSNAISYALYKAFSTSEQDKDNWNGQLKLLLEWNQLDLANDEIFTND
RRWESADLQEVMFTALIKDRPKFVRLFLENGLNLRKFLTHDVLTELFSNHFSTLVYRNLQIAKNSYNDALLTFVWKLVAN
FRRGFRKEDRNGRDEMDIELHDVSPITRHPLQALFIWAILQNKKELSKVIWEQTRGCTLAALGASKLLKTLAKVKNDINA
AGESEELANEYETRAVELFTECYSSDEDLAEQLLVYSCEAWGGSNCLELAVEATDQHFIAQPGVQNFLSKQWYGEISRDT
KNWKIILCLFIIPLVGCGFVSFRKKPVDKHKKLLWYYVAFFTSPFVVFSWNVVFYIAFLLLFAYVLLMDFHSVPHPPELV
LYSLVFVLFCDEVRQWYVNGVNYFTDLWNVMDTLGLFYFIAGIVFRLHSSNKSSLYSGRVIFCLDYIIFTLRLIHIFTVS
RNLGPKIIMLQRMLIDVFFFLFLFAVWMVAFGVARQGILRQNEQRWRWIFRSVIYEPYLAMFGQVPSDVDGTTYDFAHCT
FTGNESKPLCVELDEHNLPRFPEWITIPLVCIYMLSTNILLVNLLVAMFGYTVGTVQENNDQVWKFQRYFLVQEYCSRLN
IPFPFIVFAYFYMVVKKCFKCCCKEKNMESSVCCFKNEDNETLAWEGVMKENYLVKINTKANDTSEEMRHRFRQLDTKLN
DLKGLLKEIANKIK
```

Figs. 12A and 12B

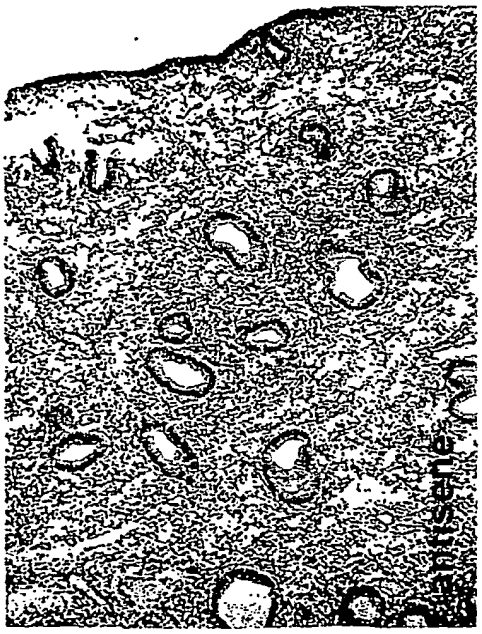




Figs. 12C and 12D

Endometrium:

E



F

Figs. 12E and 12F

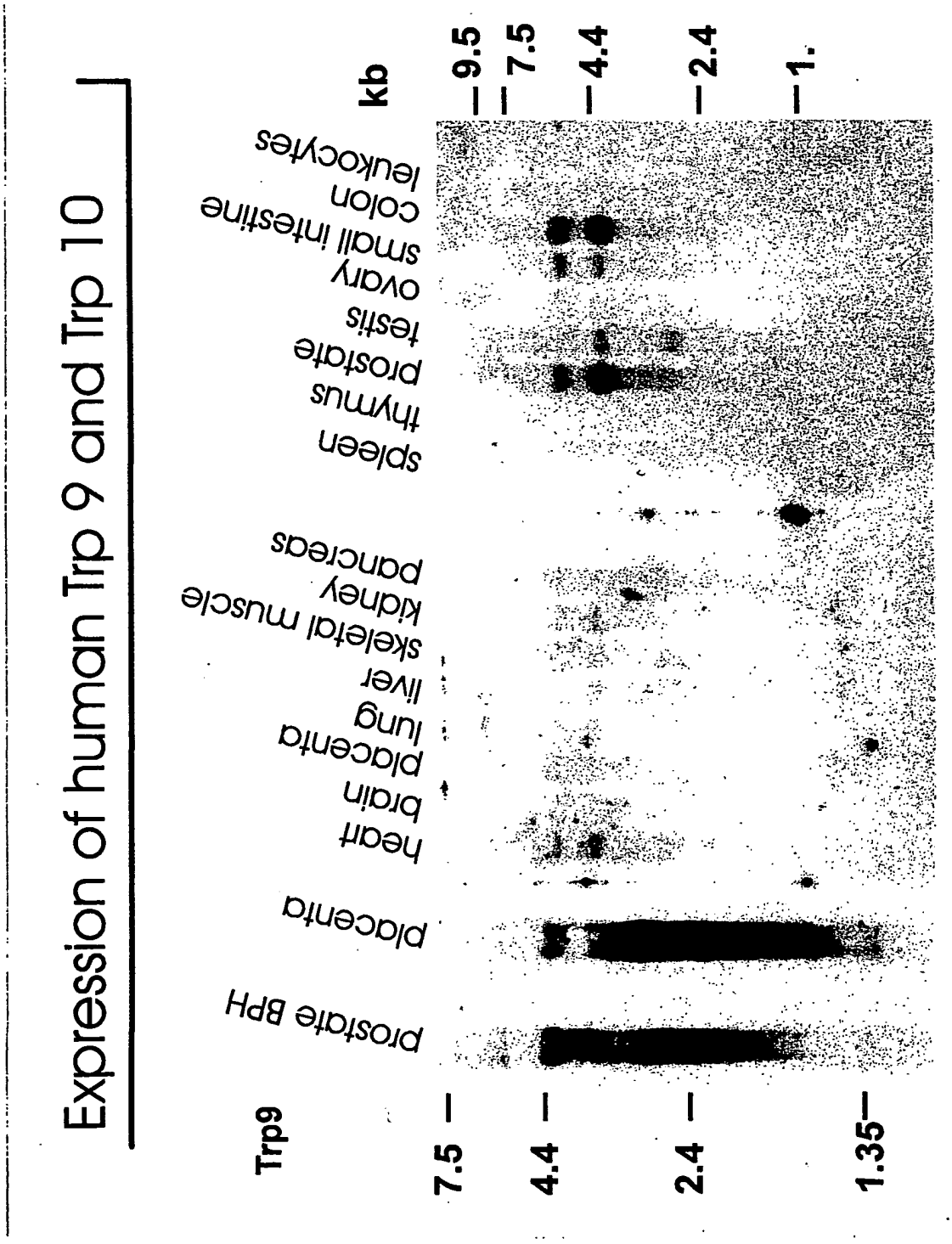

Expression of human Trp 9 and Trp 10
37/42
kb
9.5
7.5
4.4
2.4
1.
leukocytes
colon
small intestine
ovary
testis
prostate
thymus
spleen
pancreas
kidney
skeletal muscle
liver
lung
placenta
brain
heart
placenta
prostate BPH
Trp9
7.5
4.4
2.4
1.35

Fig. 13

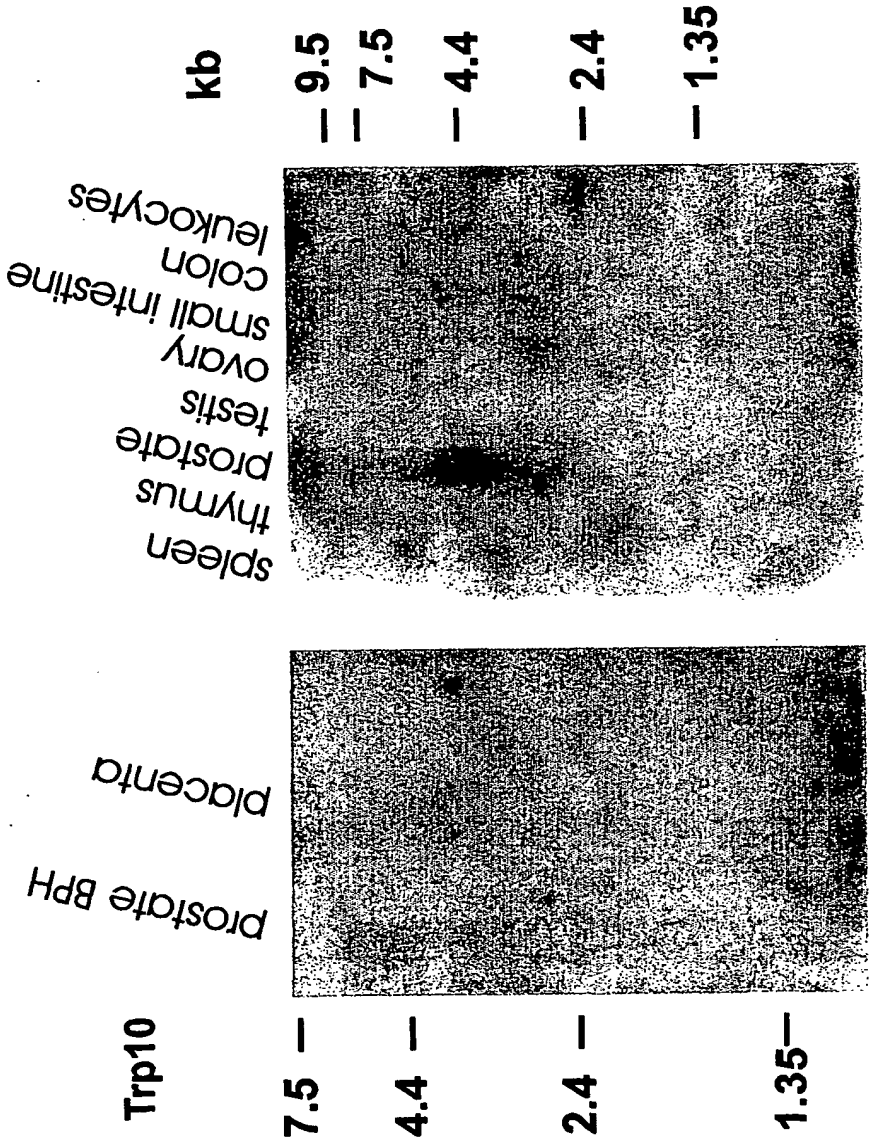
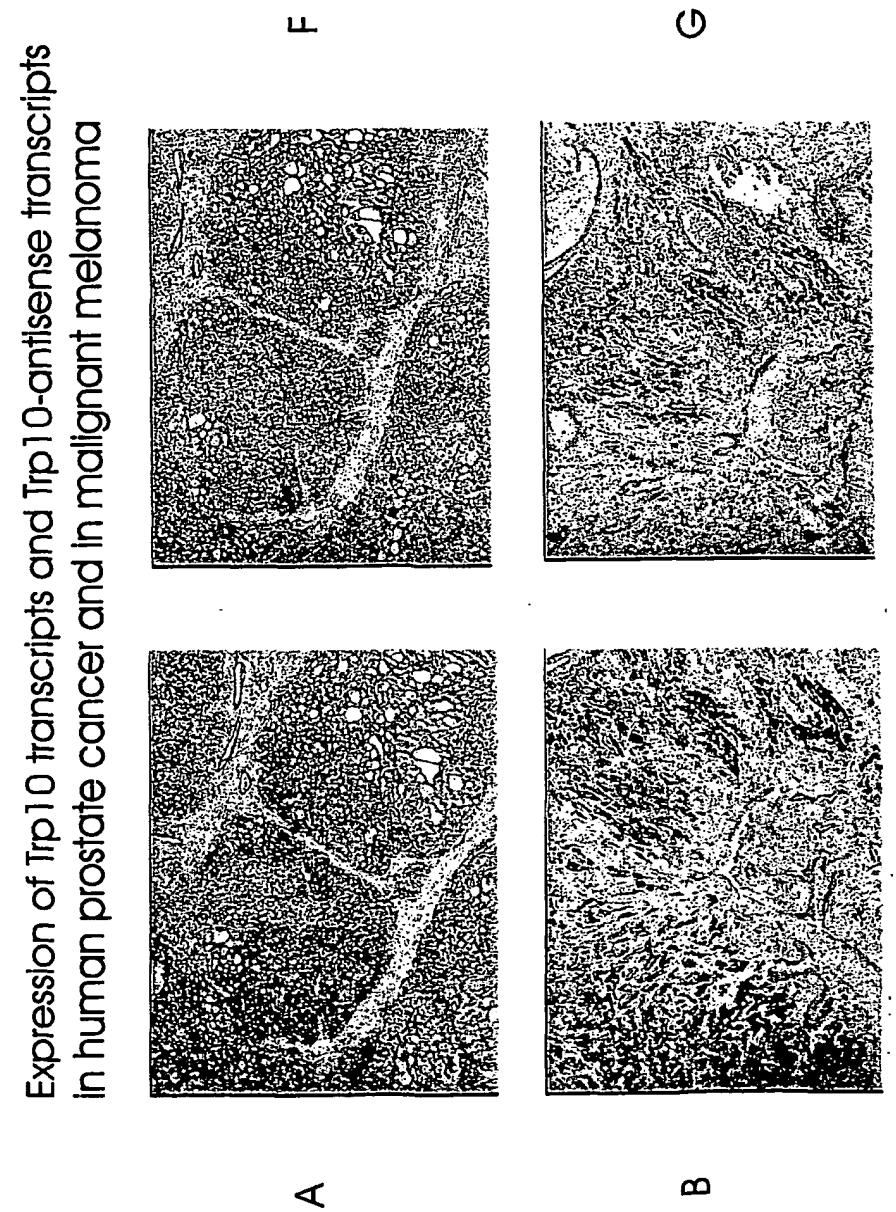
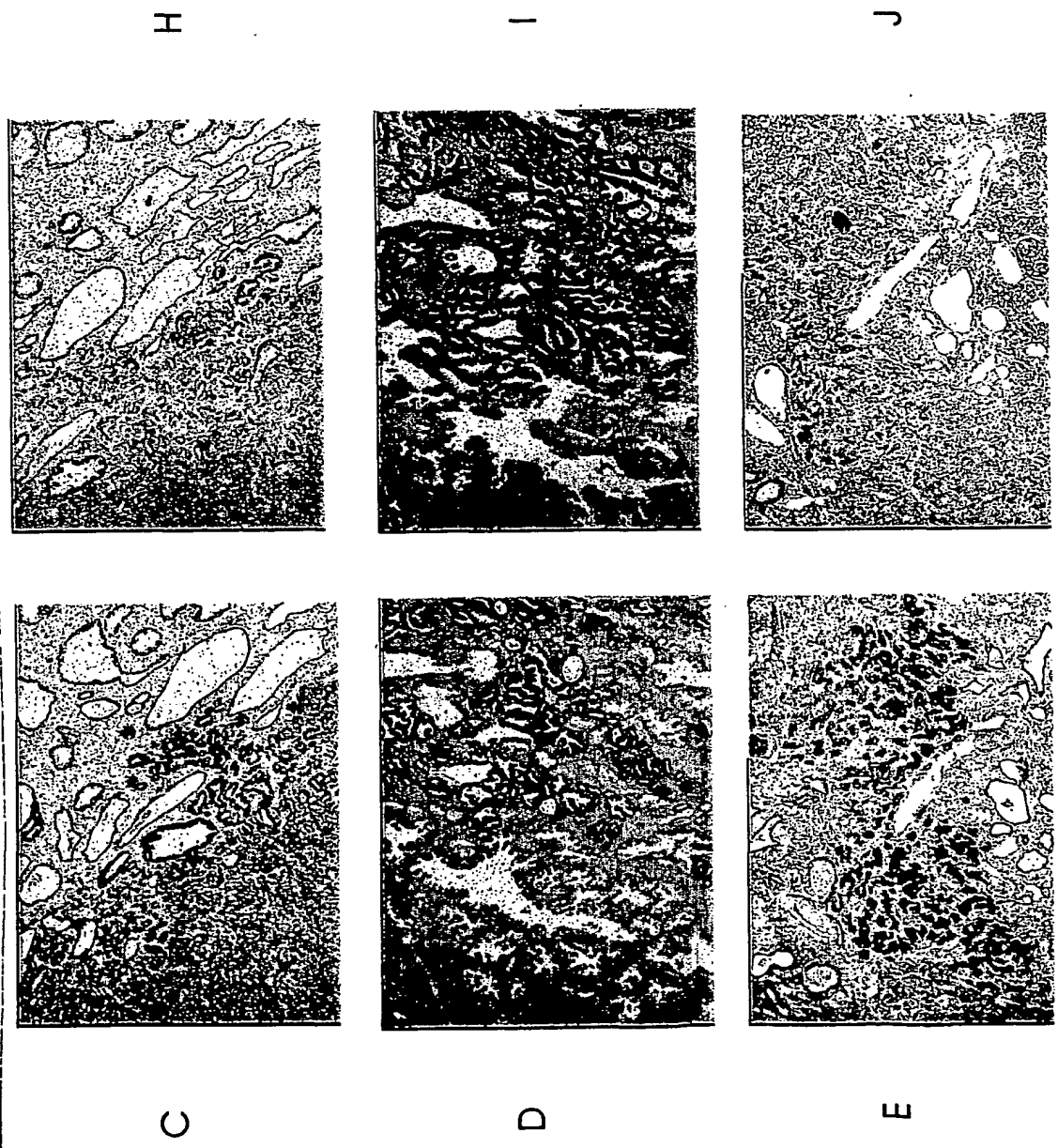


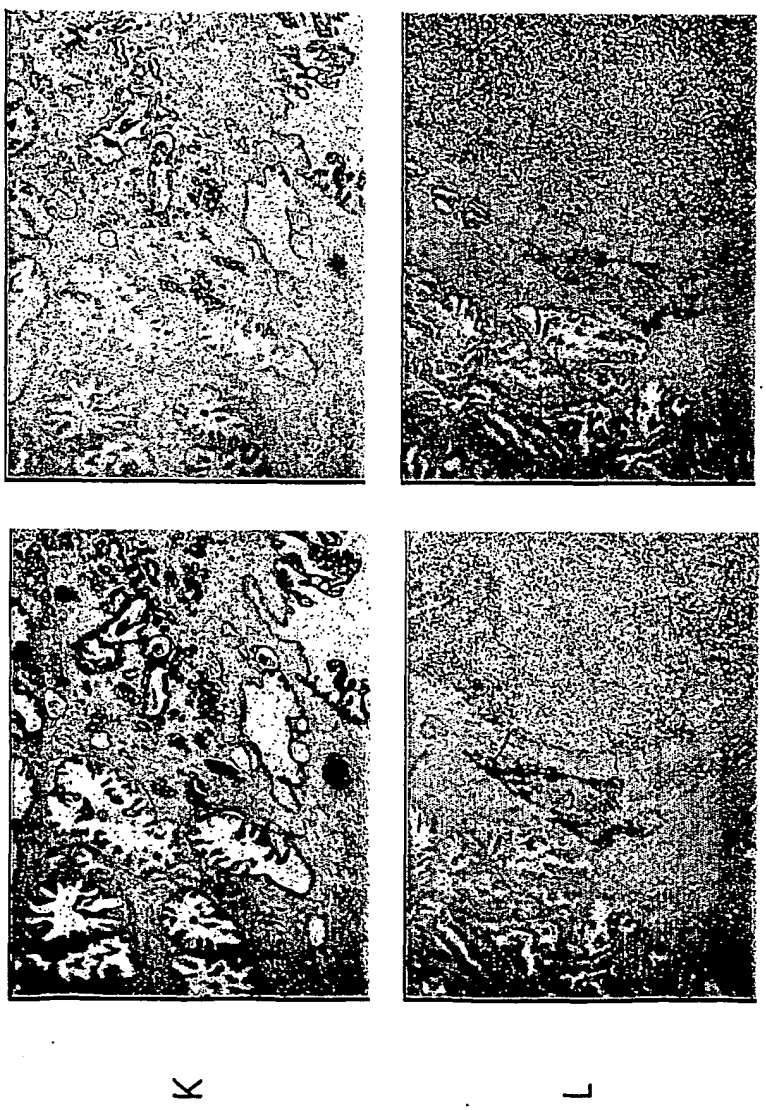
Fig. 13 / Continuation 1

Figs. 14A, 14B, 14F and 14G



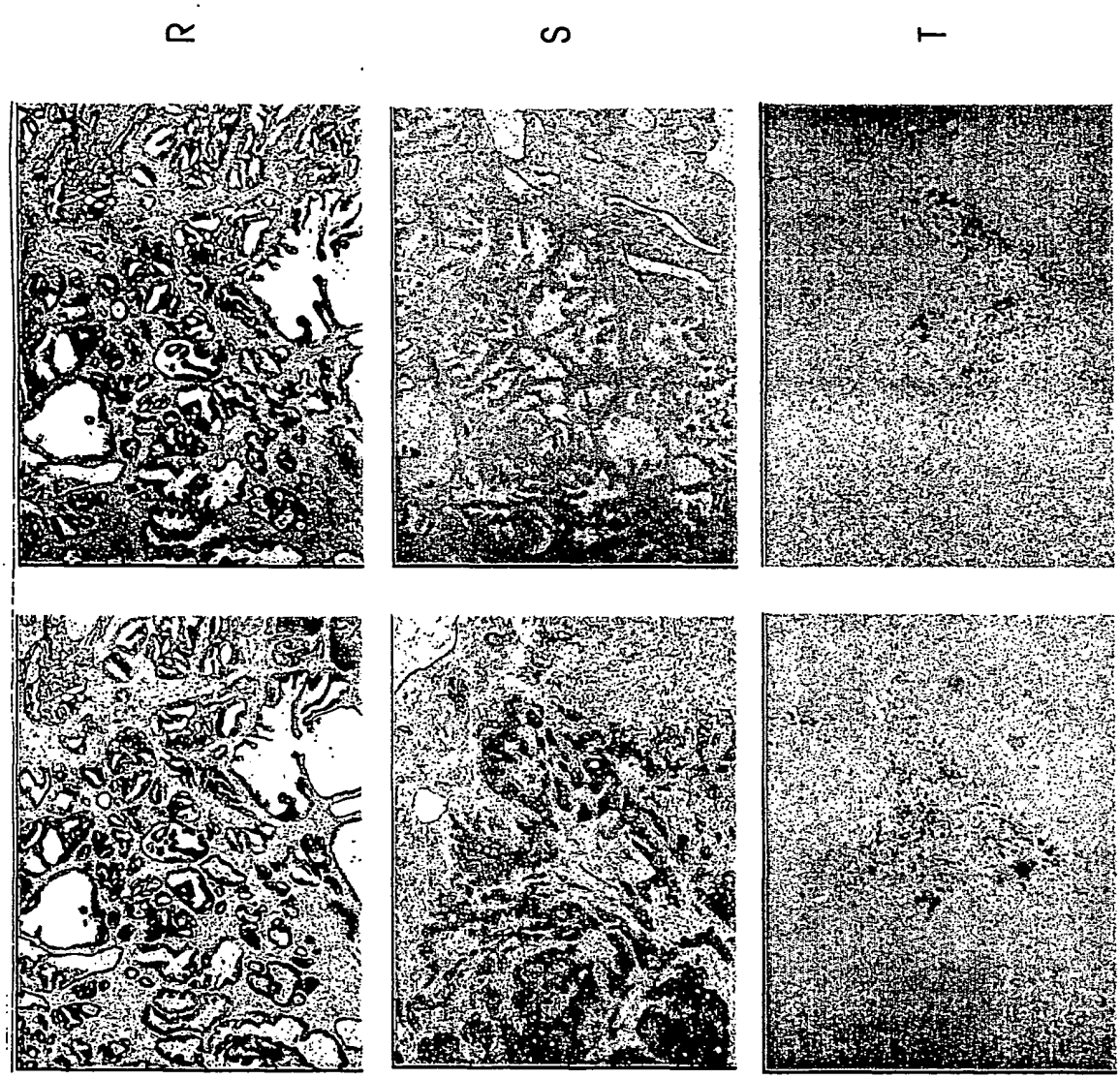


Figs. 14C, 14D, 14E, 14H, 14I and 14J



Figs. 14K, 14L, 14P and 14Q

41/42



Figs. 14M, 14N, 14O, 14R, 14S and 14T